

Copyright GenCore version 5.1.6 (c) 1993 - 2005 Compugen Ltd.

Om nucleic - nucleic search, using BW model

Run on: November 15, 2005, 23:53:58 ; Search time 31202 Seconds (without alignments)

11774.484 Million cell updates/sec

Title: US-10-717-580-11
Perfect score: 7582
Sequence: 1 caacaatcgggataaaacc..... tattaaatcttgcarctgcr 7582

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Databse :

```

1: "gb_ba:*
2: "gb_htg:*
3: "gb_in:*
4: "gb_om:*
5: "gb_ov:*
6: "gb_pat:*
7: "gb_ph:*
8: "gb_pl:*
9: "gb_pr:*
10: "gb_pro:*
11: "gb_sbs:*
12: "gb_sv:*
13: "gb_un:*
14: "gb_vti:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	7489.5	98.8	7582	BD196245 Endogenat
2	7489.6	98.8	7582	BD196245 Endogenat
3	7489.6	98.8	7582	AX00966 Sequence
4	7489.5	98.8	7582	AX027480 Sequence
5	6448	85.0	10222	AY101582 Homo sapi
6	6448	85.0	10222	AY101585 Homo sapi
7	6448	85.0	56093	AX229572 Sequence
8	6448	85.0	14194	AC000064 Human BAC
9	6448	85.0	14194	AC007566 Homo sapi
10	6446.4	85.0	10222	AY101583 Homo sapi
11	6440	84.9	10222	AY101584 Homo sapi
12	6436	84.9	10499	BD22108 Nucleic s
13	6436	84.9	10499	AX007980 Sequence
14	6346.4	83.7	10229	AY101586 Pan trogl
15	6344.8	83.7	10229	AY101587 Pan trogl
16	6344.8	83.7	18675	AC145964 Pan trogl
17	6269.6	82.7	10230	AY101588 Gorilla g
18	6269.6	82.7	10230	AY101589 Gorilla g
19	6037.8	79.6	10122	AY101590 Pongo pyg

ALIGNMENTS

Qy	2161	CTTCCTGTCGGACAACTGTCCAGATCTCACTTATCTGAGGGGTCTNTAAGCC	Qy	3241	GCRTARYAGGYTCTCGCGAWATGATCCCCAGGTWGGCRAATAACCGCCYATTA	3300
Db	2161	CTTCCTGTCGGACAACTGTCCAGATCTCACTTATCTGAGGGGTCTNTAAGCC	Db	3241	GCRTARYAGGYTCTCGCGAWATGATCCCCAGGTWGGCRAATAACCGCCYATTA	3300
Qy	2221	GCGACTCACTAGATACTTTTCCAGGCCACTAGTATGACTGGGAGCTTATCTT	Qy	3301	WATCAGSTAATTAGGAACCTAGAAGCCATACCAATTARTAGATGAYAAGTGAA	3360
Db	2221	GCGACTCACTAGATACTTTTCCAGGCCACTAGTATGACTGGGAGCTTATCTT	Db	3301	WATCAGSTAATTAGGAACCTAGAAGCCATACCAATTARTAGATGAYAAGTGAA	3360
Qy	2281	TCACTATGCTTTCTAATTAGCTGAAAGGCCACTACTTGTAGGAGACATTCA	Qy	3361	GYMRAGTGCCTTCAGGCCCTAAAGAGGGAAACCCGTTGTTAGGAGACATCA	3420
Db	2281	TCACTATGCTTTCTAATTAGCTGAAAGGCCACTACTTGTAGGAGACATCA	Db	3361	GYMRAGTGCCTTCAGGCCCTAAAGAGGGAAACCCGTTGTTAGGAGACATCA	3420
Qy	2341	GCAAAGCUGGGCATATACACCTGAACTAGGAGAGAACCCGTTGTTGNNCC	Qy	3481	TGCCAACRGGCAAGACTTSTTYATYRTCACAGAAAACAGRAAYGCTTRGGA	3480
Db	2341	GCAAAGCUGGGCATATACACCTGAACTAGGAGAGAACCCGTTGTTGNNCC	Db	3481	TGCCAACRGGCAAGACTTSTTYATYRTCACAGAAAACAGRAAYGCTTRGGA	3480
Qy	2401	CCTGTTGAGGAAGGATTAACTCTGAACTAGGAGAGAACCCGTTGTTGNNCC	Qy	3481	TGCCAACRGGCAAGACTTSTTYATYRTCACAGAAAACAGRAAYGCTTRGGA	3480
Db	2401	CCTGTTGAGGAAGGATTAACTCTGAACTAGGAGAGAACCCGTTGTTGNNCC	Db	3481	TGCCAACRGGCAAGACTTSTTYATYRTCACAGAAAACAGRAAYGCTTRGGA	3480
Qy	2461	CAAAAATGCCGTCCTGTCAGTAAGGATTAAGGATTCACCTCTTCCCTACAA	Qy	3541	GANGTAGGCAAGGGTGRCTCATGTTAGGGTAGTGGGGAGCTAGTGGYKTA	3600
Db	2461	CAAAAATGCCGTCCTGTCAGTAAGGATTAAGGATTCACCTCTTCCCTACAA	Db	3541	GANGTAGGCAAGGGTGRCTCATGTTAGGGTAGTGGGGAGCTAGTGGYKTA	3600
Qy	2521	GGCGTAGTACCCCTAGACCCAGGCCAAQAGGATTCAAAAGATGTTAGACTAA	Qy	3601	GTATCTGAGCAAGCTTCTTGTGACCTGACACVYRGCVACTGASTAGGAAT	3540
Db	2521	GGCGTAGTACCCCTAGACCCAGGCCAAQAGGATTCAAAAGATGTTAGACTAA	Db	3601	GTATCTGAGCAAGCTTCTTGTGACCTGACACVYRGCVACTGASTAGGAAT	3540
Qy	2581	AAGGCAAAAGCTAGTAACCAACGATCACTCCGTGCTTAATCCGTAGTGTAG	Qy	3661	GTAAVYRGCACTACTGCTAAGGGAGCTTGTGGTAGCTGGTCAACAGYTAA	3660
Db	2581	AAGGCAAAAGCTAGTAACCAACGATCACTCCGTGCTTAATCCGTAGTGTAG	Db	3661	GTAAVYRGCACTACTGCTAAGGGAGCTTGTGGTAGCTGGTCAACAGYTAA	3660
Qy	2641	GAGGCAACGAAACCCAGTGAGCACTGGGGTTAGTGCAGAGTTCGATTCAATG	Qy	3721	TRTCAGSCTTAACTCTGCTAAGGGAGCTTGTGGTAGCTGGTCAACAGYTAA	3720
Db	2641	GAGGCAACGAAACCCAGTGAGCACTGGGGTTAGTGCAGAGTTCGATTCAATG	Db	3721	TRTCAGSCTTAACTCTGCTAAGGGAGCTTGTGGTAGCTGGTCAACAGYTAA	3720
Qy	2701	ASGCCGTTGTTCTTATACCCAGCTGACTACCCCTTAATCTGTTGTTCCAAATAC	Qy	3781	CCAGYCNCAATTCTCCAGACATGAGAAAGATAAYAATCTCAACARTAATT	3840
Db	2701	ASGCCGTTGTTCTTATACCCAGCTGACTACCCCTTAATCTGTTGTTCCAAATAC	Db	3781	CCAGYCNCAATTCTCCAGACATGAGAAAGATAAYAATCTCAACARTAATT	3840
Qy	2761	CAAGGAGGAGGAGGTGTTACASTCTCGGACTTCTGAGTACCTTCTGATCCG	Qy	3841	TCTCAAAACCTATGCCACTCGAGGGAGACTTGTAGARTTCTCYTTGACTGATCCYGAATT	3900
Db	2761	CAAGGAGGAGGAGGTGTTACASTCTCGGACTTCTGAGTACCTTCTGATCCG	Db	3841	TCTCAAAACCTATGCCACTCGAGGGAGACTTGTAGARTTCTCYTTGACTGATCCYGAATT	3900
Qy	2821	TACATCTGCACTCTCAATTCTCTGCTTGTGAGTACCTCAACCCAREATCTCAC	Qy	3901	CAACTCTTAACTCTGAGGAGCTTGTGAGAAGAGACTCTGAGAAAGYGGGTITGC	3960
Db	2821	TACATCTGCACTCTCAATTCTCTGCTTGTGAGTACCTCAACCCAREATCTCAC	Db	3901	CAACTCTTAACTCTGAGGAGCTTGTGAGAAGAGACTCTGAGAAAGYGGGTITGC	3960
Qy	2881	TACACTGGACTATTTACCCAGGGTTAGYCCCTATTTGGCAGGAT	Qy	3961	AGTGGTAGTGTATGTAATYTGAGATATCCCTACTCCAGGAACACTAGTGTYA	4020
Db	2881	TACACTGGACTATTTACCCAGGGTTAGYCCCTATTTGGCAGGAT	Db	3961	AGTGGTAGTGTATGTAATYTGAGATATCCCTACTCCAGGAACACTAGTGTYA	4020
Qy	2941	TAGCCCGAGCTTGAGCTGACCTTCACCTGACACTCTGCTCTGCTGTAGTGTAG	Qy	4021	GCTRGGAGACTAATGCCCTCYTAKGGGACTAGTAAATTAGGAGAAGRAAAAGGGAA	4080
Db	2941	TAGCCCGAGCTTGAGCTGACCTTCACCTGACACTCTGCTCTGCTGTAGTGTAG	Db	4021	GCTRGGAGACTAATGCCCTCYTAKGGGACTAGTAAATTAGGAGAAGRAAAAGGGAA	4080
Qy	3001	ATTACCTTTCYCYCRTCTAGAAACCTTGCCATCAAGCACCACAGRCCTCTMA	Qy	4081	TATATACAGCTCTTAACTGCTTGTACTAGTCAGTCAGTCAGTCAGTCAGTCAGTC	4140
Db	3001	ATTACCTTTCYCYCRTCTAGAAACCTTGCCATCAAGCACCACAGRCCTCTMA	Db	4081	TATATACAGCTCTTAACTGCTTGTACTAGTCAGTCAGTCAGTCAGTCAGTCAGTC	4140
Qy	3061	ATTCTCTGCACTCTGCTGCTACCTGAGGATCAAAACSSARACTCARCTCTCACGC	Qy	4141	AGAAGGAAATCTCTRACTTCYGAAGAACCTTCAACAGGAGCTATGCGTCA	4200
Db	3061	ATTCTCTGCACTCTGCTGCTACCTGAGGATCAAAACSSARACTCARCTCTCACGC	Db	4141	AGAAGGAAATCTCTRACTTCYGAAGAACCTTCAACAGGAGCTATGCGTCA	4200
Qy	3121	AGGTTAAATTAGCTGCTATCTCATCYCAAACCTTAACGACTAAGRRTCTG	Qy	4201	ATTATATGCGTAAAGAACCTTCAACAGGAGCTATGCGTCA	4260
Db	3121	AGGTTAAATTAGCTGCTATCTCATCYCAAACCTTAACGACTAAGRRTCTG	Db	4201	ATTATATGCGTAAAGAACCTTCAACAGGAGCTATGCGTCA	4260
Qy	3181	TCCAGCTATCTGCTTATCTCATCYCAAACCTTAACGACTAAGRRTCTG	Qy	4261	NAAGGAAGRAAGGAAATASAAGRGAATGCGTAAAGCMAAAGAGC	4320
Db	3181	TCCAGCTATCTGCTTATCTCATCYCAAACCTTAACGACTAAGRRTCTG	Db	4261	NAAGGAAGRAAGGAAATASAAGRGAATGCGTAAAGCMAAAGAGC	4320
Qy	3240		Qy	4321	TGCAAGGAGGACCCTCATTAGAAATGCTTATTAACTTCCCTTAGTATAGGGTATCC	4380

Db	4321	TGCAGGCGAGACCTCCATATTAGATAACTTCCTTGTAGTTAGGGTAATCC	4380	Qy
Qy	4381	CTTCCGGAAACCGGCCAGACTCAGGGAACAGATGGGAACTCACGAGG	4440	Db
Db	4381	CTTCCGGAAACCGGCCAGACTCAGGGAACAGATGGGAACTCACGAGG	4440	Qy
Qy	4441	CAGTTTCTCCCCTGGGAGGGTTAGCCACTGAGAAGGAATACTTGCTGCA	4500	Db
Db	4441	CAGTTTCTCCCCTGGGAGGGTTAGCCACTGAGAAGGAATACTTGCTGCA	4500	Qy
Qy	4501	TATCCCATGGAATTACTTAACCTCTCATCACACTTCACTTAGCATGAG	4560	Db
Db	4501	TATCCCATGGAATTACTTAACCTCTCATCACACTTCACTTAGCATGAG	4560	Qy
Qy	4561	CCATCARATGCCAATCATATTACTGACCAGGCCAGTACCGAGGAA	4620	Db
Db	4561	CCATCARATGCCAATCATATTACTGACCAGGCCAGTACCGAGGAA	4620	Qy
Qy	4621	AKTCAGGCGCTGTGAAKTOGTGCAACTTCAACATGAGAAGGAATAC	4680	Db
Db	4621	AKTCAGGCGCTGTGAAKTOGTGCAACTTCAACATGAGAAGGAATAC	4680	Qy
Qy	4681	AGGARAAACAAACAGGCATTACCTGACRARAARACTGCACTGATT	4740	Db
Db	4681	AGGARAAACAAACAGGCATTACCTGACRARAARACTGCACTGATT	4740	Qy
Qy	4741	CCCACACTCAGGGATTCACTTACTGTTAGGTTACCAACAG	4800	Db
Db	4741	CCCACACTCAGGGATTCACTTACTGTTAGGTTACCAACAG	4800	Qy
Qy	4801	RAGGCCTTCCCCGTGAGAACAGGAAAGGCTTAAAGGCTAAGGCTA	4860	Db
Db	4801	RAGGCCTTCCCCGTGAGAACAGGAAAGGCTTAAAGGCTAAGGCTA	4860	Qy
Qy	4861	ATATTCAGATTGGACTTCCGGAGGTACAGTGACATAGCCTGTTCCAG	4920	Db
Db	4861	ATATTCAGATTGGACTTCCGGAGGTACAGTGACATAGCCTGTTCCAG	4920	Qy
Qy	4921	GCCACAGTAACCCAGGGATATCCAGGGTTAGGATACTACATGCGCC	4980	Db
Db	4921	GCCACAGTAACCCAGGGATATCCAGGGTTAGGATACTACATGCGCC	4980	Qy
Qy	4981	TGAGGCCACAGTCTCAGGGAGGTGAGAATGAAATGAAATGAAATGAA	5040	Db
Db	4981	TGAGGCCACAGTCTCAGGGAGGTGAGAATGAAATGAAATGAAATGAA	5040	Qy
Qy	5041	AAAGCAACCGAACCCACCTCACATGGCTGTCGTGCTAGCTAA	5100	Db
Db	5041	AAAGCAACCGAACCCACCTCACATGGCTGTCGTGCTAGCTAA	5100	Qy
Qy	5101	AGATCTGCAACTTCCCAAAGCAGGACTTAGCCATACGAAATGCTGTA	5160	Db
Db	5101	AGATCTGCAACTTCCCAAAGCAGGACTTAGCCATACGAAATGCTGTA	5160	Qy
Qy	5161	CCCTCATACCATGACCTGGCTGACCCAGACGCCACTAGTGCA	5220	Db
Db	5161	CCCTCATACCATGACCTGGCTGACCCAGACGCCACTAGTGCA	5220	Qy
Qy	5221	CCTCCATGCAATATCAACAGTTCTAAACATTACAGAACCTATCC	5280	Db
Db	5221	CCTCCATGCAATATCAACAGTTCTAAACATTACAGAACCTATCC	5280	Qy
Qy	5281	CCCGATCCCTAGTACCTCTGGAGGACCTACCGTCAATCTGAGAG	5340	Db
Db	5281	CCCGATCCCTAGTACCTCTGGAGGACCTACCGTCAATCTGAGAG	5340	Qy
Qy	5341	AGGAAAGAAACTATTCACCTCTGGAGGACCTACCGTCAATCTGAGAG	5400	Db
Db	5341	AGGAAAGAAACTATTCACCTCTGGAGGACCTACCGTCAATCTGAGAG	5400	Qy
Qy	5401	CGGTTAAAGTGGCTGGAGTTCTGGATACATCACAGGAACTTAC	5460	Db
Db	5401	CGGTTAAAGTGGCTGGAGTTCTGGATACATCACAGGAACTTAC	5460	Qy
Qy	5461	TGCACAAAGAACCTGAAATTCGGAGAACAGCTAGCTATCTGTGAACTCTAGAGG	5520	Db
Db	5461	TGCACAAAGAACCTGAAATTCGGAGAACAGCTAGCTATCTGTGAACTCTAGAGG	5520	Qy
Qy	5521	ATTGCGCTGCTTCACACACACAGGAGGAAGTAATAATTC	5580	Db
Db	5521	ATTGCGCTGCTTCACACACACAGGAGGAAGTAATAATTC	5580	Qy
Qy	5581	ATGSCCCTCCCTCATATTTCCTCTTGTGTTTTCACCTCTACT	5640	Db
Db	5581	ATGSCCCTCCCTCATATTTCCTCTTGTGTTTTCACCTCTACT	5640	Qy
Qy	5641	GCACCCCMCCAGCCGCGCTGATGAGCTGCTCCCTYACMAGAGTTCTTGAGA	5700	Db
Db	5641	GCACCCCMCCAGCCGCGCTGATGAGCTGCTCCCTYACMAGAGTTCTTGAGA	5700	Qy
Qy	5701	ATGAGCGTCCCGAAATATGATGCCCATGTTAGGAGCTTSTAGGAAACCC	5760	Db
Db	5701	ATGAGCGTCCCGAAATATGATGCCCATGTTAGGAGCTTSTAGGAAACCC	5760	Qy
Qy	5761	ACCTTCACTGCCCCAACCTTATGCCGCACTGCTATCACCTGCCACTCTTGATG	5820	Db
Db	5761	ACCTTCACTGCCCCAACCTTATGCCGCACTGCTATCACCTGCCACTCTTGATG	5820	Qy
Qy	5821	CATGCAAACTACTCATATTGGACGAGAAATGATAATCTGTGCTTGTGAGGACT	5880	Db
Db	5821	CATGCAAACTACTCATATTGGACGAGAAATGATAATCTGTGCTTGTGAGGACT	5880	Qy
Qy	5881	GGAGTCACTGTGTTGAGCTACTTCACCCAACTGCTATGATGGGGTGGAGT	5940	Db
Db	5881	GGAGTCACTGTGTTGAGCTACTTCACCCAACTGCTATGATGGGGTGGAGT	5940	Qy
Qy	5941	CGAGATCAGGAGAGAAACAGTGAATAGGAAATCTCCAACTCACCGGGTACAT	6000	Db
Db	5941	CGAGATCAGGAGAGAAACAGTGAATAGGAAATCTCCAACTCACCGGGTACAT	6000	Qy
Qy	6001	GCACCTCTAGCCCTACAAACGACTAGTCTCAAACTACATGAAACCCCGCGTAC	6060	Db
Db	6001	GCACCTCTAGCCCTACAAACGACTAGTCTCAAACTACATGAAACCCCGCGTAC	6060	Qy
Qy	6061	CATACTGGCTGTAAGCTTAATACACCTCTACTGGCTCATGGCTCGGCC	6120	Db
Db	6061	CATACTGGCTGTAAGCTTAATACACCTCTACTGGCTCATGGCTCGGCC	6120	Qy
Qy	6121	CAAAACCTACTACTGTGGATGTGCTCCCTGACTTCAGCCATATGTTCA	6180	Db
Db	6121	CAAAACCTACTACTGTGGATGTGCTCCCTGACTTCAGCCATATGTTCA	6180	Qy
Qy	6181	CCTGACTCTGACATGAACTCAGCAGAAATTAACCACTTGTGTTAGTA	6240	Db
Db	6181	CCTGACTCTGACATGAACTCAGCAGAAATTAACCACTTGTGTTAGTA	6240	Qy
Qy	6241	GAACCTCTGTCTCAATSTGGAAATAACCCATACTCAAACTAACCTGTTGAAATT	6300	Db
Db	6241	GAACCTCTGTCTCAATSTGGAAATAACCCATACTCAAACTAACCTGTTGAAATT	6300	Qy
Qy	6301	ACGRATACTACATACACCAACTCCCATGTTAGGTTACTCTCCACACAA	6360	Db
Db	6301	ACGRATACTACATACACCAACTCCCATGTTAGGTTACTCTCCACACAA	6360	Qy
Qy	6361	ATAGCTCTCTGTTCAATSTGGAAATAACCACTAACCTGTTGTTAGT	6420	Db
Db	6361	ATAGCTCTCTGTTCAATSTGGAAATAACCACTAACCTGTTGTTAGT	6420	Qy
Qy	6421	ATGGCTCTGAGATCTGAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6480	Db
Db	6421	ATGGCTCTGAGATCTGAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6480	Qy
Qy	6481	ATGCAACAGATATAGTATGTTAGCTTAACTGCGCAACAGAGTACCT	6540	Db
Db	6481	ATGCAACAGATATAGTATGTTAGCTTAACTGCGCAACAGAGTACCT	6540	Qy

QY	6541	CCTCCCTTTGTTAGGAGCAGGAGTGTAGGTGCACTTAGGTACTGSCATCGGGTAC	6600	RESULT 2	B0267487	BD267487	7582 bp	DNA	linear	PAT 17-JUL-2003
Db	6541	CTTCCTTTGTTAGGAGCAGGAGTGTAGGTGCACTTAGGTACTGSCATCGGGTAC	6600	LOCUS						
QY	6601	ACRACCTCTACTCAGTTCTACAACTCTTCAGAACTTAATGGGCATGGAAAGG	6660	DEFINITION						
Db	6601	ACRACCTCTACTCAGTTCTACAACTCTTCAGAACTTAATGGGCATGGAAAGG	6660	Endogenous nucleic fragment						
QY	6661	GTCGCCGACTCCCTGGTACCTGCAAGATACTTAACTTAACTCCCTAGCAGCTAGTC	6720	marking method and reagent.						
Db	6661	GTCGCCGACTCCCTGGTACCTGCAAGATACTTAACTTAACTCCCTAGCAGCTAGTC	6720	VERSTON	B0267487	JP 2002314980-A/28.				
QY	6721	CTAACATGGAAGAGCTTGTAGCTGCTAACCTCTCAAGACTAAATGGGCATGGAAAGG	6780	KEYWORDS						
Db	6721	CTAACATGGAAGAGCTTGTAGCTGCTAACCTCTCAAGACTAAATGGGCATGGAAAGG	6780	SOURCE						
QY	6781	GGGGAGAATGTTGTTATTGTTAATCACTTAACTGGGATCTGACTGAGAAGTTRAGAA	6840	ORGANISM						
Db	6781	GGGGAGAATGTTGTTATTGTTAATCACTTAACTGGGATCTGACTGAGAAGTTRAGAA	6840	Homo sapiens (human)						
QY	6841	ATTCAGATGAAATCACGTTAGCAGGAGCTGAGTCAAGACTGGACCTGGGCTC	6900	COMMENT						
Db	6841	ATTCAGATGAAATCACGTTAGCAGGAGCTGAGTCAAGACTGGACCTGGGCTC	6900	JOURNAL						
QY	6901	CTCAGGCRATGATGCCCTGTGATTCTCCCTCTTACGAGCTTACGAGCTTAAATG	6960	PP						
Db	6901	CTCAGGCRATGATGCCCTGTGATTCTCCCTCTTACGAGCTTACGAGCTTAAATG	6960	DD	JP-2002534980-A/28	JP 2002534980-A/28				
QY	6961	CTACTCTCTTGACCTGTACTCTTACCTCTTACCTCTTACCTCTTACCTCTTAC	7020	PR						
Db	6961	CTACTCTCTTGACCTGTACTCTTACCTCTTACCTCTTACCTCTTACCTCTTAC	7020	PI						
QY	7021	GAAGCTTAACTACAATGGAGCCAGTCAGTCAAGACTAGTAACTTACCGAGA	7080	GLAUCIA PARAHOS BACCALA, FRANCOIS MALLET, CECILE VOISSET PC						
Db	7021	GAAGCTTAACTACAATGGAGCCAGTCAGTCAAGACTAGTAACTTACCGAGA	7080	G01N3/569, G01N3/569, G01N3/58//C12B1/02, C12B15/00 CC						
QY	7081	CCCTGGACCGCCCTGTAGCCACGGACTGTTAGTAAAGGACCCCTCT	7140	PC						
Db	7081	CCCTGGACCGCCCTGTAGCCACGGACTGTTAGTAAAGGACCCCTCT	7140	PC						
QY	7141	GAGGAATCTCTAGCTCACACTCTACTAGCCCATTTAGCAGGAGGTAGC	7200	ENDOGENETIC NUCLEIC						
Db	7141	GAGGAATCTCTAGCTCACACTCTACTAGCCCATTTAGCAGGAGGTAGC	7200	NUCLEIC FRAGMENT						
QY	7201	GGTGGCGCCACCTCCCAACAGCACTGTTCTGTAGATGGGGACTGAGA	7260	disease,						
Db	7201	GGTGGCGCCACCTCCCAACAGCACTGTTCTGTAGATGGGGACTGAGA	7260	marking method and reagent						
QY	7261	GCAGGACTAGCTGTGATTCTCTAGGTGAYTAAAGTCTYAGCTAGCTAGTGGAGGT	7320	CC						
Db	7261	GCAGGACTAGCTGTGATTCTCTAGGTGAYTAAAGTCTYAGCTAGCTAGTGGAGGT	7320	FH						
QY	7321	ACCACATCCACTTTAACACGGGGCTGCACTTGTACCTGACCACTACAGAG	7380	FT						
Db	7321	ACCACATCCACTTTAACACGGGGCTGCACTTGTACCTGACCACTACAGAG	7380	source						
QY	7381	CTCACTAACTTAATGGCAAGAGCAGGGTAAAGAAATAGCAATCATYATG	7440	ORIGIN						
Db	7381	CTCACTAACTTAATGGCAAGAGCAGGGTAAAGAAATAGCAATCATYATG	7440	Query Match						
QY	7440	CTCACTAACTTAATGGCAAGAGCAGGGTAAAGAAATAGCAATCATYATG	7440	98.8%; Score: 749.6; DB: 6; Length: 7582;						
Db	7440	CTCACTAACTTAATGGCAAGAGCAGGGTAAAGAAATAGCAATCATYATG	7440	Best Local Similarity	100.0%; Pred. No.: 0;					
QY	7501	GGGCAACCCCTTGGTCCCTCCCTGGTACGGCTTCACTGCTGAGCTTCTG	7560	Matches: 7582; Conservative	0; Mismatches: 0; Indels: 0; Gaps: 0;					
Db	7501	GGGCAACCCCTTGGTCCCTCCCTGGTACGGCTTCACTGCTGAGCTTCTG	7560	Key						
QY	7561	TCTATAAATGTCARCTGC	7582	1. 7582						
Db	7561	TCTATAAATGTCARCTGC	7582	/organism='Homo sapiens'						
				/mol_type="genomic DNA"						
				/db_xref="taxon:3606"						

QY	421	TGGCCCAAGATTCATTCTTGGAAATCGTAGGSCAACGAACTCCAGGTAGAGATA	480	Db	1501	ACACAGGGAGGAAGAAATCTRACTGCCTTCTGAGAAGACTRAGGGAGGATTAG	1560
Db	421	TGGCCCAAGATTCATTCTTGGAAATCGTAGGSCAACGAACTCCAGGTAGAGATA	480	QY	1561	GAAGCCTGCCCTCTGTCACCTGACTCTTCCTGAGGCCAACTAATCTTAAGGTAGTT	1620
QY	481	GARGCTTGCCACCATCTTGGAGGCGGCTGTCACCRCTTGGAGGTTGACACATC	540	Db	1561	GAAGCCTGCCCTCTGTCACCTGACTCTTCCTGAGGCCAACTAATCTTAAGGTAGTT	1620
Db	481	GARGCTTGCCACCATCTTGGAGGCGGCTGTCACCRCTTGGAGGTTGACACATC	540	QY	1621	TATCATCTAGCTAGTCAGACATGGAAAGAAAACCTCCAAAGTCTGGCTAGGGCGAG	1680
QY	541	TGGGAGCTCTGGAGGAGGCCCCCGGATACATTGGGRACCMRSAGGACATCC	600	Db	1621	TATCATCTAGCTAGTCAGACATGGAAAGAAAACCTCCAAAGTCTGGCTAGGGCGAG	1680
Db	541	TGGGAGCTCTGGAGGAGGCCCCCGGATACATTGGGRACCMRSAGGACATCC	600	QY	1681	CAAACCTAGAAACCCATTGACTCTGGCAACYTCCGTTTATATAGATCAGG	1740
QY	601	MAGTGTAGGGAAAGTTCGCCAACAGCACGCCCCTAGACGTTGGRACCMRSAGGACATCC	660	Db	1681	CAAACCTAGAAACCCATTGACTCTGGCAACYTCCGTTTATATAGATCAGG	1740
Db	601	MAGTGTAGGGAAAGTTCGCCAACAGCACGCCCCTAGACGTTGGRACCMRSAGGACATCC	660	QY	1741	GGGAGGGAGACGGCAACGGGAAAGGGATAAAAAAAGCCACOGCTTGTAGCTGACCT	1800
QY	661	TGGGACATTGACCTCAGACACTAAGAACGACCTTATTTCTCAGTCTGACCT	720	Db	1741	GGGAGGGAGACGGCAACGGGAAAGGGATAAAAAAAGCCACOGCTTGTAGCTGACCT	1800
Db	661	TGGGACATTGACCTCAGACACTAAGAACGACCTTATTTCTCAGTCTGACCT	720	QY	1801	CAGGAGTGGACTTGGAGGCTCTGGAAAGGAAGTCTGGCAATATGATGCTAA	1860
QY	721	GGCTGGCACTCTGGGGAGGAGATAANTTAAACACATCTACAGCTAGCYCTTTG	780	Db	1801	CAGGAGTGGACTTGGAGGCTCTGGAAAGGAAAGTCAGGAAATGATGCTGCTAA	1860
Db	721	GGCTGGCACTCTGGGGAGGAGATAANTTAAACACATCTACAGCTAGCYCTTTG	780	QY	1861	TAGGCTGTCTCCAGGGCTTACAGGACATTAAAGATGTCAGTAG	1920
QY	781	TAGAAAGGCATAAGGTGAGTGCTAGTCATACTTCTTGTAGAGACAA	840	Db	1861	TAGGCTGTCTCCAGGGCTTACAGGACATTAAAGATGTCAGTAG	1920
Db	781	TAGAAAGGCATAAGGTGAGTGCTAGTCATACTTCTTGTAGAGACAA	840	QY	1921	TRAGCCGCCCTCTGRCATGCCCCATTTCAGGAACTTAAAGGAACTCTGGAGGCCACTGCC	1980
QY	841	CYCAACAATTGTTAAAGGTGATTTGGCTTACAGGAACTTAAAGGAACTCTGGAGGCCACTGCC	900	Db	1921	TRAGCCGCCCTCTGRCATGCCCCATTTCAGGAACTTAAAGGAACTCTGGAGGCCACTGCC	1980
Db	841	CYCAACAATTGTTAAAGGTGATTTGGCTTACAGGAACTTAAAGGAACTCTGGAGGCCACTGCC	900	QY	1981	CAGGAGCAAGCTTGTGACTCAGGCGCTACAGGAACTTAAAGGAACTCTGGAGGCCACTGCC	2040
QY	901	CTATCCAGCATCCCGACTCTTCCCMYTAATAAGGACCCCTTCACCCAATGG	960	Db	1981	CAGGAGCAAGCTTGTGACTCAGGCGCTACAGGAACTTAAAGGAACTCTGGAGGCCACTGCC	2040
Db	901	CTATCCAGCATCCCGACTCTTCCCMYTAATAAGGACCCCTTCACCCAATGG	960	QY	2041	AGGTGCTGGGCAAGGGCATCCATGCCCACCTCACAGGCTGSGTATGCT	2100
QY	961	TCCAAAGGAGATAGACAAAGGTAACGGTACCCATATTCCCAT	1020	Db	2041	AGGTGCTGGGCAAGGGCATCCATGCCCACCTCACAGGCTGSGTATGCT	2100
Db	961	TCCAAAGGAGATAGACAAAGGTAACGGTACCCATATTCCCAT	1020	QY	2101	GGCATCTAGGGCCAGGAGGAACTTGTGACTCTGGAACTCTGGCTGCGTCCTCTAGCTACT	2160
QY	1021	TATGACCCCTCCAGCTGGAGGAGAATGGCCAGCAGTCAGTCTGCT	1080	Db	2101	GGCATCTAGGGCCAGGAGGAACTTGTGACTCTGGAACTCTGGCTGCGTCCTCTAGCTACT	2160
Db	1021	TATGACCCCTCCAGCTGGAGGAGAATGGCCAGCAGTCAGTCTGCT	1080	QY	2161	CTCTGCGCCGACACTGTCCTCCAGAACTCTGACTCTTCAGGGGTCCNTAAAGACG	2220
QY	1081	TTYYTCCTCCAGACTTAAGCAAAACAGACTTGTGAAATTCTGAGTAACTYCT	1140	Db	2161	CTCTGCGCCGACACTGTCCTCCAGAACTCTGACTCTTCAGGGGTCCNTAAAGACG	2220
Db	1081	TTYYTCCTCCAGACTTAAGCAAAACAGACTTGTGAAATTCTGAGTAACTYCT	1140	QY	2221	GGCGTCACTAGATACTTYTCAGGCACTAGTTGAACTCTGGAGGCTTCTCT	2280
QY	1141	GATGGCTATATGTTGAGGTTAACGAACTTGTGAAATTCTGAGTAACTYCT	1200	Db	2221	GGCGTCACTAGATACTTYTCAGGCACTAGTTGAACTCTGGAGGCTTCTCT	2280
Db	1141	GATGGCTATATGTTGAGGTTAACGAACTTGTGAAATTCTGAGTAACTYCT	1200	QY	2281	TCACATGCTTCTAAATTGCTGAAAGCCCACACTACCTTGTGAGGAGACATCTA	2340
QY	1201	ATATATGCTCTGTAATAGACACTAACCCAAATGAGAACTGCCACATAC	1260	Db	2281	TCACATGCTTCTAAATTGCTGAAAGCCCACACTACCTTGTGAGGAGACATCTA	2340
Db	1201	ATATATGCTCTGTAATAGACACTAACCCAAATGAGAACTGCCACATAC	1260	QY	2341	GCAGAACGCGGGCCATATACCTGACATAGGAGGAAACCCGGTGTGTGNCC	2400
QY	1261	AGCTGAGGTTTGGGACTCTCTGGATCTCAGTGTGAAATTCTGAGTAACTYCT	1320	Db	2341	GCAGAACGCGGGCCATATACCTGACATAGGAGGAAACCCGGTGTGTGNCC	2400
Db	1261	AGCTGAGGTTTGGGACTCTCTGGATCTCAGTGTGAAATTCTGAGTAACTYCT	1320	QY	2401	CCTCTGAGGAGGAAATTCTGAACTCTGGAGGAAACAGAGGAACATATGGAGGAC	2460
QY	1321	GAAGGAAGGAAATGATTCCTCACGGGCGCAGCTTCCAGCTTACCTCAGT	1380	Db	2401	CCTGCTGAGGAGGAAATTCTGAACTCTGGAGGAAACAGAGGAACATATGGAGGAC	2460
Db	1321	GAAGGAAGGAAATGATTCCTCACGGGCGCAGCTTCCAGCTTACCTCAGT	1380	QY	2461	CAAGAGTGGCCCTCTGTCAGTAACCTAAGGATCCACTCTTCCCTACCAA	2520
QY	1381	GGGAGCAAGAAATGATTCCTCACGGGCGCAGCTTCCAGCTTACCTCAGT	1440	Db	2461	CAAGAGTGGCCCTCTGTCAGTAACCTAAGGATCCACTCTTCCCTACCAA	2520
Db	1381	GGGAGCAAGAAATGATTCCTCACGGGCGCAGCTTCCAGCTTACCTCAGT	1440	QY	2521	GGCAGTACCCCTCAGACCCAGGCCAACAGGATTCAGGATGTTAGGACTAA	2580
QY	1441	TASAGGACTTGGGAAACTCTAGTAACTCTAGTGTGTCACCTGTCACCTA	1500	Db	2521	GGCAGTACCCCTCAGACCCAGGCCAACAGGATTCAGGATGTTAGGACTAA	2580
Db	1441	TASAGGACTTGGGAAACTCTAGTAACTCTAGTGTGTCACCTGTCACCTA	1500	QY	2581	AAGCCCAAGCTTGTAGTAAACCTACTGCCTTCCTGAGAGACTAAGGGAGGATTAG	2640
QY	1501	ACACAGGGAGGAAGAAATCTRACTGCCTTCTGAGAAGACTRAGGGAGGATTAG	1560				

Qy	3721	TRTCAGGCTTAACTTGATGARGGCCAGTGTGCRACATGTCACACTCTCAC 3780
Db	3721	TRTCAGGCTTAACTTGATGARGGCCAGTGTGCRACATGTCACACTCTCAC 3780
Qy	3781	CCAGYNCATTTCTCAGACATGAGAAGATAAYAATAGTCACARTAAT 3840
Db	3781	CCAGYNCATTTCTCAGACATGAGAAGATAAYAATAGTCACARTAAT 3840
Qy	3841	TCTCAACACTATGCCACTCGGGACCTTGTAGGTTAGGTTCCYTACTGTCACCT 3900
Db	3841	TCTCAACACTATGCCACTCGGGACCTTGTAGGTTAGGTTCCYTACTGTCACCT 3900
Qy	3901	CAACTGTAACTGTGGAATCTTGTGAGAATCTCAACCCARCATCTAAC 2880
Db	3901	CAACTGTAACTGTGGAATCTTGTGAGAATCTCAACCCARCATCTAAC 2880
Qy	3960	CAACTGTAACTGTGGAATCTTGTGAGAATCTCAACCCARCATCTAAC 2880
Db	3960	CAACTGTAACTGTGGAATCTTGTGAGAATCTCAACCCARCATCTAAC 2880
Qy	3961	AGTGGTAGGTGATAATGGAATAYTGAAGTAATCCCTCATCCAGGACTAGTCYACT 4020
Db	3961	AGTGGTAGGTGATAATGGAATAYTGAAGTAATCCCTCATCCAGGACTAGTCYACT 4020
Qy	4021	GCTRCAGAACTAATAGCTTCATYKGGGACTAGATTAGGAGAAGRAAAAGGAA 4080
Db	4021	GCTRCAGAACTAATAGCTTCATYKGGGACTAGATTAGGAGAAGRAAAAGGAA 4080
Qy	4081	TATATATACGACTCTRARRATGCTTACCTGTCCTGCATGCCATMRGCAATATSAR 4140
Db	4081	TATATATACGACTCTRARRATGCTTACCTGTCCTGCATGCCATMRGCAATATSAR 4140
Qy	4141	AGAAAGGAAATCTTAACCTCYGAGRAACCTATCACACATCAGGAGCATTAGAR 4200
Db	4141	AGAAAGGAAATCTTAACCTCYGAGRAACCTATCACACATCAGGAGCATTAGAR 4200
Qy	4201	ATTATAYTGGCGWGTACAGAACTTGTGAGAATCTGCGTCTGAGGAA 4260
Db	4201	ATTATAYTGGCGWGTACAGAACTTGTGAGAATCTGCGTCTGAGGAA 4260
Qy	4261	NAANGGAAGGAAAGGAAATASAGRGAAATGGCAAGCKATATGAGCMAAAGAG 4320
Db	4261	NAANGGAAGGAAAGGAAATASAGRGAAATGGCAAGCKATATGAGCMAAAGAG 4320
Qy	4321	TGCAGGCGACCCCTCCATAGAAATGCTTATAAACTCCCTAGTAGGTAATC 4380
Db	4321	TGCAGGCGACCCCTCCATAGAAATGCTTATAAACTCCCTAGTAGGTAATC 4380
Qy	4380	CTTCGGGAAACCAAGCCCGACTCTAGAGGAAACAGAATGGGAACCTCAGAGG 4440
Db	4381	CTTCGGGAAACCAAGCCCGACTCTAGAGGAAACAGAATGGGAACCTCAGAGG 4440
Qy	4441	CAGTTTCCCTCGGAGCGGTAGCCACTGTGAGAGGGAAATACTTTCCTGCAA 4500
Db	4441	CAGTTTCCCTCGGAGCGGTAGCCACTGTGAGAGGGAAATACTTTCCTGCAA 4500
Qy	4501	TATCCATGGAAATTACTTAAACCCCTCATCAAACCTTCACTTAGGATGATGCA 4560
Db	4501	TATCCATGGAAATTACTTAAACCCCTCATCAAACCTTCACTTAGGATGATGCA 4560
Qy	4561	CCATCARATGCCAAATCATTTACTGACCGACGCCCTTCACAACTACAGCAR 4620
Db	4561	CCATCARATGCCAAATCATTTACTGACCGACGCCCTTCACAACTACAGCAR 4620
Qy	4621	AKTCAGGGCCGTGCAKGTCAGTGCACARAATACTCCCTCCCTATGCCAGCTCTC 4680
Db	4621	AKTCAGGGCCGTGCAKGTCAGTGCACARAATACTCCCTCCCTATGCCAGCTCTC 4680
Qy	4681	AGGARAACAAARAACAGGCCATTACCTGGRARAARACTGGCAACTGATTACCCAG 4740
Db	4681	AGGARAACAAARAACAGGCCATTACCTGGRARAARACTGGCAACTGATTACCCAG 4740
Qy	4740	4741 CCCAACCTCGGAGATTTCAGTATCTAGTCTCGGGTAACTTCAGGGTGGCA 4800
Db	4741 CCCAACCTCGGAGATTTCAGTATCTAGTCTCGGGTAACTTCAGGGTGGCA 4800	

QY	4801	RAGGCCTTCCCTGAGGACAGAAAGGCCAAGGGTAATAAGCACTAGTTCTGAA	4860	Db
QY	4801	RAGGCCTTCCCTGAGGACAGAAAGGCCAAGGGTAATAAGCACTAGTTCTGAA	4860	Db
QY	4861	ATATTCCAGATGGACCTCCCGGAGGTTAGAGGACATACTGCCTGTCTCCG	4920	Db
QY	4861	ATATTCCAGATGGACCTCCCGGAGGTTAGAGGACATACTGCCTGTCTCCG	4920	Db
QY	4921	GCCAGAGTACCCAGGAGTCCAGGCTTACCTACACTGCCCC	4980	Db
QY	4921	GCCAGAGTACCCAGGAGTCCAGGCTTACCTACACTGCCCC	4980	Db
QY	4981	TGAGGCCACAGTCAGGGAGTCCAGGCTTACCTACACTGCCCC	4980	Db
QY	4981	TGAGGCCACAGTCAGGGAGTCCAGGCTTACCTACACTGCCCC	4980	Db
QY	5041	AAARGAACCCGGAAACCCACTCACATGGCTGCTGTCTATAGCTTAAA	5100	Db
QY	5041	AAARGAACCCGGAAACCCACTCACATGGCTGCTGTCTATAGCTTAAA	5100	Db
QY	5101	AGAACTGCACTTCCCAGAAAGCAGGACTAGCCATAGGAAATGTA	5160	Db
QY	5101	AGAACTGCACTTCCCAGAAAGCAGGACTAGCCATAGGAAATGTA	5160	Db
QY	5161	CCCTCTATACCAATGACCTTGTTGCTGACCCAACTGAGGATGAGG	5160	Db
QY	5161	CCCTCTATACCAATGACCTTGTTGCTGACCCAACTGAGGATGAGG	5160	Db
QY	5221	CCTCTTAGCAAATACAAAGTCTTAAACCTACAGGAACTTCCCTGAGG	5220	Db
QY	5221	CCTCTTAGCAAATACAAAGTCTTAAACCTACAGGAACTTCCCTGAGG	5220	Db
QY	5221	CCTCTTAGCAAATACAAAGTCTTAAACCTACAGGAACTTCCCTGAGG	5220	Db
QY	5281	AGGGAAAGAACTATCCACCCWNGACATGGTTAGTCAGTCCTCYCTTA	5340	Db
QY	5281	AGGGAAAGAACTATCCACCCWNGACATGGTTAGTCAGTCCTCYCTTA	5340	Db
QY	5341	CCCTCTCTAGATCATCTGGGAGGACCTACCCAGCTATTATVACCCACTG	5400	Db
QY	5341	CCCTCTCTAGATCATCTGGGAGGACCTACCCAGCTATTATVACCCACTG	5400	Db
QY	5401	CGGTAAAGGGCTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5460	Db
QY	5401	CGGTAAAGGGCTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5460	Db
QY	5461	TGCCAAAGGAACTGAAATCAGGACAAAGCTAGCTTCTGTGAACTCTAGG	5520	Db
QY	5461	TGCCAAAGGAACTGAAATCAGGACAAAGCTAGCTTCTGTGAACTCTAGG	5520	Db
QY	5521	ATTCGGCTGCTCTCTAACACACCAACGGAGGAAGTAACTAAATCTAAATCCCC	5580	Db
QY	5521	ATTCGGCTGCTCTCTAACACACCAACGGAGGAAGTAACTAAATCTAAATCCCC	5580	Db
QY	5581	ATGGSCTCTTACATATTCTCTKASTGSTTYYACCTSTTCACTCTACT	5640	Db
QY	5581	ATGGSCTCTTACATATTCTCTKASTGSTTYYACCTSTTCACTCTACT	5640	Db
QY	5641	GCACCCCTCATGGCTGTATGACCGAGTACTCCCTAACMAGATTCTATGAGA	5700	Db
QY	5641	GCACCCCTCATGGCTGTATGACCGAGTACTCCCTAACMAGATTCTATGAGA	5700	Db
QY	5701	ATGCAGCTCTGGAAATATGATGCCCATGTTAGGACTCTTSTAAGGAACCCC	5760	Db
QY	5701	ATGCAGCTCTGGAAATATGATGCCCATGTTAGGACTCTTSTAAGGAACCCC	5760	Db
QY	5761	ACCTCACTGCCAACCCATATGCCGCACTCTATCTCTCCACCTTGTCT	5820	Db
QY	5761	ACCTCACTGCCAACCCATATGCCGCACTCTATCTCTCCACCTTGTCT	5820	Db
QY	5821	CATGAACTCTATGGACAGGAAATGATTAATCTCTAGTGTCTGAGGAGACT	5880	Db
QY	5821	CATGAACTCTATGGACAGGAAATGATTAATCTCTAGTGTCTGAGGAGACT	5880	Db
QY	5881	GGAGTCACTCTGTGGACTACTTCACCAACATGCTGAGTGTCTGAGGAGACT	5940	Db
QY	5881	GGAGTCACTCTGTGGACTACTTCACCAACATGCTGAGTGTCTGAGGAGACT	5940	Db
QY	5941	CAAGATCGGAGAGAAAGTAAAGAGTAATCTCCAACTCAGGGTAC	6000	Db
QY	5941	CAAGATCGGAGAGAAAGTAAAGAGTAATCTCCAACTCAGGGTAC	6000	Db
QY	6001	GGCACTCTAGGCCCTAAAGAACTAGCTCTAAACTATGAAACCTCGGCC	6060	Db
QY	6001	GGCACTCTAGGCCCTAAAGAACTAGCTCTAAACTATGAAACCTCGGCC	6060	Db
QY	6061	CAPACTCCTGTAACCTTAACTTAACTACCTACTGGCTCATGGCTCGGCC	6120	Db
QY	6061	CAPACTCCTGTAACCTTAACTTAACTACCTACTGGCTCATGGCTCGGCC	6120	Db
QY	6121	CAAACCTACTACTGTGATGCTCCCTGACTTARGCCATATGTTCACT	6180	Db
QY	6121	CAAACCTACTACTGTGATGCTCCCTGACTTARGCCATATGTTCACT	6180	Db
QY	6181	CTCTACTCTGAACTGAACTCTGAGCTCTGGCTCATGGCTCGGCC	6240	Db
QY	6181	CTCTACTCTGAACTGAACTCTGAGCTCTGGCTCATGGCTCGGCC	6240	Db
QY	6241	GAACCTCTGTCTCACTSTGAAATAACCCATACCTCAACCTACCTGTTAAATT	6300	Db
QY	6241	GAACCTCTGTCTCACTSTGAAATAACCCATACCTCAACCTACCTGTTAAATT	6300	Db
QY	6301	AGCAATACATACACACCAACTCCCATGATGCTGGTAACTCCACACAA	6360	Db
QY	6301	AGCAATACATACACACCAACTCCCATGATGCTGGTAACTCCACACAA	6360	Db
QY	6361	ATAGCTCTCTACCTACCTCGGAAATTTCGTCGTCGACTCTGGTTAGA	6420	Db
QY	6361	ATAGCTCTCTACCTACCTCGGAAATTTCGTCGTCGACTCTGGTTAGA	6420	Db
QY	6421	AATGGCTCTCTGAACTCTGAACTCTGAGCTCTCTCTGAGCCCTATGRCATCTAC	6480	Db
QY	6421	AATGGCTCTCTGAACTCTGAGCTCTGAGCTCTCTCTGAGCCCTATGRCATCTAC	6480	Db
QY	6481	ACTGACAGAGTTATAGCTGTTAGTGTCTATAGCCGCGACAAAGAGTACCT	6540	Db
QY	6481	ACTGACAGAGTTATAGCTGTTAGTGTCTATAGCCGCGACAAAGAGTACCT	6540	Db
QY	6541	CTTCCTTTGTTAGGAGGAGGAGGCTAGTGACACTAGGACTAGGGCTTGGGTAC	6600	Db
QY	6541	CTTCCTTTGTTAGGAGGAGGAGGCTAGTGACACTAGGACTAGGGTAC	6600	Db
QY	6601	ACAACTCTACTGCTACTACAACCTATCTAAAGAACTAAATGGGAATCTGAGG	6660	Db
QY	6601	ACAACTCTACTGCTACTACAACCTATCTAAAGAACTAAATGGGAATCTGAGG	6660	Db
QY	6661	GTCCGGGACTCCCTGGTCACTTCAGGAACTCTAACCTCTAGGAGCTGAGCTCT	6720	Db
QY	6661	GTCCGGGACTCCCTGGTCACTTCAGGAACTCTAACCTCTAGGAGCTGAGCTCT	6720	Db
QY	6721	CTAACATGAGAGCTTAACTCTGAGGAGGAGGAGCTTATTTTA	6780	Db
QY	6721	CTAACATGAGAGCTTAACTCTGAGGAGGAGGAGCTTATTTTA	6780	Db
QY	6781	GGGGAGAGATGCTGTTATGTTATCTACCTCCGGATGCTACTGAGAAGTGTAAAGA	6840	Db
QY	6781	GGGGAGAGATGCTGTTATGTTATCTACCTCCGGATGCTACTGAGAAGTGTAAAGA	6840	Db
QY	6841	ATTCGAGCTCTGGAAATATGATGCCCATGTTAGGACTCTTSTAAGGAACCCC	6900	Db
QY	6841	ATTCGAGCTCTGGAAATATGATGCCCATGTTAGGACTCTTSTAAGGAACCCC	6900	Db
QY	6901	CTACCTCTTGTGAGGCTCTTCTAGGAGCTCTAGAGCTTACCTGCTCTCCAGATC	7020	Db
QY	6901	CTACCTCTTGTGAGGCTCTTCTAGGAGCTCTAGAGCTTACCTGCTCTCCAGATC	7020	Db
QY	6961	CTACCTCTTGTGAGGCTCTTCTAGGAGCTCTAGAGCTTACCTGCTCTCCAGATC	7020	Db

Db 6961 CTACTCTCTTGGACCTGTATCTRACTCTTGTAACCTGTCCTGCAAGAATC 7020 Qy 61 CTCCTCTTGATGGGAGTGTTCATGCTATTCACTTAATCTGCACATGCA 120 Db 7021 GAAGCTGTRAACTCAAGGAGCCAGATCGTCCAGAATGAGTCTACGGAGA 7080 Qy 61 CTCCTCTTGATGGGAGTGTTCATGCTATTCACTTAATCTGCACATGCA 120 Db 7081 CCCCTGGACGGCCTGTYAACGATCTGATGTTATCATCAAGGCCCTCT 7140 Qy 61 CTCCTCTTGATGGGAGTGTTCATGCTATTCACTTAATCTGCACATGCA 180 Db 7081 CCCCTGGACGGCCTGTYAACGATCTGATGTTATCATCAAGGCCCTCT 7140 Qy 181 TGTGTGCCCCACCGCAGAACCTCGCTGACTCCATCTGATGTTATCATCAAGGCCCTCT 7140 Db 7141 GAGGAATTCTAGTGACACCTCTACTAGTGCCCAATTCAGCAGAACGAGTAGAC 7200 Qy 181 TGTGTGCCCCACCGCAGAACCTCGCTGACTCCATCTGATGTTATCATCAAGGCCCTCT 7200 Db 7141 GAGGAATTCTAGTGACACCTCTACTAGTGCCCAATTCAGCAGAACGAGTAGAC 7200 Qy 181 TGTGTGCCCCACCGCAGAACCTCGCTGACTCCATCTGATGTTATCATCAAGGCCCTCT 7200 Db 7201 GGTSGTGCGCCAACTCCCAAACAGCACTTAGGTTCCTGTTGAGATGGGACTGTGA 7260 Qy 181 TGTGTGCCCCACCGCAGAACCTCGCTGACTCCATCTGATGTTATCATCAAGGCCCTCT 7260 Db 7201 GGTSGTGCGCCAACTCCCAAACAGCACTTAGGTTCCTGTTGAGATGGGACTGTGA 7260 Qy 181 TGTGTGCCCCACCGCAGAACCTCGCTGACTCCATCTGATGTTATCATCAAGGCCCTCT 7260 Db 7261 GACAGGAATGAGTCTAGGTGATGTTAGCTAGCTGAATGAGATGGGACTGTGA 7320 Qy 181 TGTGTGCCCCACCGCAGAACCTCGCTGACTCCATCTGATGTTATCATCAAGGCCCTCT 7320 Db 7261 GACAGGAATGAGTCTAGCTAGCTGAATGAGATGGGACTGTGA 7320 Qy 181 TGTGTGCCCCACCGCAGAACCTCGCTGACTCCATCTGATGTTATCATCAAGGCCCTCT 7320 Db 7321 ACCACATCCACTTAAACAGGGGCTTCAACTAGYVACACTGACCATGAGAG 7380 Qy 361 GGTTCCATGGTCTCTGACCCAGGCTTAACTACATTTATGCTGAGAG 7380 Db 7321 ACCACATCCACTTAAACAGGGGCTTCAACTAGYVACACTGACCATGAGAG 7380 Qy 421 TGCCCCAGATTCATCTGTTGGAGCTCTGAGGACACTTACAGTACATTTATGCTGAGAG 7440 Db 7381 CTCACTAAATGCTTAATGGCAAGAGCAGGAGCTTAAGAAATGCCCATTATTTAC 7440 Qy 421 TGCCCCAGATTCATCTGTTGGAGCTCTGAGGACACTTACAGTACATTTATGCTGAGAG 7440 Db 7441 MTGAGACACAGGAGGACATGATGGATAACCAACTGVTGAGCCGAA 7500 Qy 481 GARGCTGCACTGGCTCTGAGGACACTTACAGTACATTTATGCTGAGAG 7500 Db 7441 MTGAGACACAGGAGGACATGATGGATAACCAACTGVTGAGCCGAA 7500 Qy 541 TTGGAGCTCTGAGGAGAACGGGCTCTGAGGACACTTACAGTACATTTATGCTGAGAG 7560 Db 7501 CGGCAACCCCTTGGTCCCTCTTGTATGGAGCTCTGTTGATGTTATTCAC 7560 Qy 601 MAAGTGTGGAAACGTTCCCGCAACACAAACCGCCCTAGACCTATTGGGRACAMSRAAGCACATCC 600 Db 7501 CGGCAACCCCTTGGTCCCTCTTGTATGGAGCTCTGTTGATGTTATTCAC 7560 Qy 601 MAAGTGTGGAAACGTTCCCGCAACACAAACCGCCCTAGACCTATTGGGRACAMSRAAGCACATCC 600 Db 7561 TCTATTAATCTGCACTGCR 7582 Qy 601 MAAGTGTGGAAACGTTCCCGCAACACAAACCGCCCTAGACCTATTGGGRACAMSRAAGCACATCC 660 Db 7561 TCTATTAATCTGCACTGCR 7582 Qy 661 TGGGAMAAATTGACCTGACACTAGAGAAGAAGGACTATATTCTCGCAGGCC 720 Db 661 TGGGAMAAATTGACCTGACACTAGAGAAGAAGGACTATATTCTCGCAGGCC 720 Qy 721 GCCTGGACTCTGAGGGAGTAAATTACACCACTTACAGTAGACYCTTTG 780 Db 721 GCCTGGACTCTGAGGGAGTAAATTACACCACTTACAGTAGACYCTTTG 780 Qy 781 TAGAAAGGCAATGGGTGAATGCTGACATAAGTACAACTTCTTTCATAGAGCA 840 Db 781 TAGAAAGGCAATGGGTGAATGCTGACATAAGTACAACTTCTTTCATAGAGCA 840 Qy 840 TAGAAAGGCAATGGGTGAATGCTGACATAAGTACAACTTCTTTCATAGAGCA 840 Db 841 CTACACATTGAAAGTGTGTTATGCTCTAGGGAGCCTTAGAGTCTACTCC 900 Qy 841 CTACACATTGAAAGTGTGTTATGCTCTAGGGAGCCTTAGAGTCTACTCC 900 Db 901 CTATCCAGATCCCCACTCTTCCACMYTAATAGGACCCCTTCACCAAAATGG 960 Qy 901 CTATCCAGATCCCCACTCTTCCACMYTAATAGGACCCCTTCACCAAAATGG 960 Db 901 CTATCCAGATCCCCACTCTTCCACMYTAATAGGACCCCTTCACCAAAATGG 960 Qy 961 TCCAAGGGAGTAGACAAAGGTAAACACTGAACCAAAGTGTGCAATTCCT 1020 Db 961 TCCAAGGGAGTAGACAAAGGTAAACACTGAACCAAAGTGTGCAATTCCT 1020 Qy 1021 TATGACCCCTCCACAGGAGTGGGAGCTGCGCAAGGAGCTGCTGCT 1080 Db 1021 TATGACCCCTCCACAGGAGTGGGAGCTGCGCAAGGAGCTGCTGCT 1080 Qy 1081 TTYTTCCTCCAGACTTAAGCAATAAACAGACTTAGTAATCTCAGATAACCT 1140 Db 1081 TTYTTCCTCCAGACTTAAGCAATAAACAGACTTAGTAATCTCAGATAACCT 1140

ORIGIN

Query Match 98.8%; Score 7489.6; DB 6; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAACATGGGGATAACCCGGCATCGAGCTGGCAACAGCAGCCCCCTTGGTCC 60
1 CAACATGGGGATAACCCGGCATCGAGCTGGCAACAGCAGCCCCCTTGGTCC 60

Oy	1141	GATGGCTATTGRTGTTACAGGGTTAGGACAATTCTCATGAGAGAT	1200	Db	2221	GGCAGTCACTAGATACTTTCRCCAGCACTAAGTGTAACTGGAGAGTTATTT	2280
Db	1141	GATGGCTATTGRTGTTACAGGGTTAGGACAATTCTCATGAGAGAT	1200	Qy	2281	TCACTATCTTCTAAATTATGCTTGAAGGCCACTRACTCTTAGGGAGACATTCA	2340
Oy	1201	ATATGTCACTGCTAAATCGACACTACCCAATGGAGATGCCACCTAATGCG	1260	Db	2281	TCACTATCTTCTAAATTATGCTTGAAGGCCACTRACTCTTAGGGAGACATTCA	2340
Db	1201	ATATGTCACTGCTAAATCGACACTACCCAATGGAGATGCCACCTAATGCG	1260	Qy	2241	GCRAAACGAGGGCCATTATACACCCTGACATAGGAGAGCACCCGTTGTCNC	2400
Oy	1261	AGCCTGAGRGTTGGCATCTGGTATCTGGTACGGTCATOGATANGATGACACA	1320	Db	2341	GCRAAACGAGGGCCATTATACACCCTGACATAGGAGAGCACCCGTTGTCNC	2400
Db	1261	AGCCTGAGRGTTGGCATCTGGTACGGTCATOGATANGATGACACA	1320	Qy	2401	CCTGCTTCAAGGAGGATTANTCCTGAGACTTGGAACAGAGGAATATGGACAGC	2460
Oy	1321	GAAGGAAGANAAATGATTCCCACAGGGCAGCARGCAGTCCAGTASACCTCATG	1380	Db	2401	CCTGCTTCAAGGAGGATTANTCCTGAGACTTGGAACAGAGGAATATGGACAGC	2460
Db	1321	GAAGGAAGANAAATGATTCCCACAGGGCAGCARGCAGTCCAGTASACCTCATG	1380	Qy	2461	CAAAAGATGCCGCTGTCAGTAACTAAGGATTCACCTCCTTCCCTACCAA	2520
Oy	1381	GGCACACAGAACTGACTACATGGAGATTTGCTGAGACATTGCTACTGTGTC	1440	Db	2461	CAAAAGATGCCGCTGTCAGTAACTAAGGATTCACCTCCTTCCCTACCAA	2520
Db	1381	GGCACACAGAACTGACTACATGGAGATTTGCTGAGACATTGCTACTGTGTC	1440	Qy	2521	GGCAGTACCCCTCAGAACCCACAGGATTAATCTGAGGATTCACAGTGTACCA	2580
Oy	1441	TAAAGGACTTAAGGAAACTGGAGAMARCTAYGAATTACTCATGATGTCACCA	1500	Db	2521	GGCAGTACCCCTCAGAACCCACAGGATTAATCTGAGGATTCACAGTGTACCA	2580
Db	1441	TAAAGGACTTAAGGAAACTGGAGAMARCTAYGAATTACTCATGATGTCACCA	1500	Qy	2581	AACCCGAGCTTAGAAACATGCTTAACCCGCTCAGTAACAGGATTCACAGTGTACCA	2640
Oy	1501	ACACAGGGGAACTGCTACTGCTACTGCTCTTGAGGAGACTAAGGAGGATTAG	1560	Db	2581	AACCCGAGCTTAGAAACATGCTTAACCCGCTCAGTAACAGGATTCACAGTGTACCA	2640
Db	1501	ACACAGGGGAACTGCTACTGCTACTGCTCTTGAGGAGACTAAGGAGGATTAG	1560	Qy	2641	GGCACACAGAAACCTGGAGCAGTGGAGGGTAGTCAGAATCTCAGGATATCATGG	2700
Oy	1561	GAAGCTCTCTGTCACTGACTGACTCTGAGCCACTAATCTAAAGTGTGCGTAGCCGAG	1620	Db	2641	GGCACACAGAAACCTGGAGCAGTGGAGGGTAGTCAGAATCTCAGGATATCATGG	2700
Db	1561	GAAGCTCTCTGTCACTGACTGACTCTGAGCCACTAATCTAAAGTGTGCGTAGCCGAG	1620	Qy	2701	AGCCCGTGTGCTTTATACCGAGCTACCTAGCTTAAACATGCTAACCTCGTACAGTATCCGTTAGTGATG	2760
Oy	1621	TATCACTGAGCAGACATTAAGAALAACTCAALAGTGTGCGTAGCCGAG	1680	Db	2701	AGCCCGTGTGCTTTATACCGAGCTACCTAGCTTAAACATGCTAACCTCGTACAGTATCCGTTAGTGATG	2760
Db	1621	TATCACTGAGCAGACATTAAGAALAACTCAALAGTGTGCGTAGCCGAG	1680	Qy	2761	CGAGGAGCAGAGCTGGAGCAGTGGAGGGTAGTCAGAATCTCAGGATATCATGG	2820
Oy	1681	CAAACTTAAACCTTATGAACTGGCACYTCYGGTTTAAATGAGTCAGAG	1740	Db	2761	CGAGGAGCAGAGCTGGAGCAGTGGAGGGTAGTCAGAATCTCAGGATATCATGG	2820
Db	1681	CAAACTTAAACCTTATGAACTGGCACYTCYGGTTTAAATGAGTCAGAG	1740	Qy	2821	TACATCTGACTCTCAATCTGTCCTTGAGAATCTGAGTACTCTAACACCACATCTAAC	2880
Oy	1741	GACGAGGCGAACGGACAAACGGGTTAAAAAAGGACCCGCTTGTAGCTGACCT	1800	Db	2821	TACATCTGACTCTCAATCTGTCCTTGAGAATCTGAGTACTCTAACACCACATCTAAC	2880
Db	1741	GACGAGGCGAACGGACAAACGGGTTAAAAAAGGACCCGCTTGTAGCTGACCT	1800	Qy	2881	TCACCTGACTCTTACCCCTAGGGTTCAGGGATTCYCCCATCTTGGCAGGAT	2940
Oy	1801	CAGGCAAGTGGACTTGGGGCTCTGGCTCTGGAAAGCTGGCAAATTGATGCTTA	1860	Db	2881	TCACCTGACTCTTACCCCTAGGGTTCAGGGATTCYCCCATCTTGGCAGGAT	2940
Db	1801	CAGGCAAGTGGACTTGGGGCTCTGGCTCTGGAAAGCTGGCAAATTGATGCTTA	1860	Qy	2941	TACCCCAGACTTGAGCARTMTCAACCTGAGACTCTGCTCTGCTGTAGTGATG	3000
Oy	1861	TAGGCTTCTCCAGTGGCTCTAACAGACCTTAAGAAGTGTGCAAGTAGAG	1920	Db	2941	TACCCCAGACTTGAGCAGTTCAGTCATCTGCTGTAGTGATG	3000
Db	1861	TAGGCTTCTCCAGTGGCTCTAACAGACCTTAAGAAGTGTGCAAGTAGAG	1920	Qy	3001	ATTACTTTRGCGCCTTCAGGAGTACTGGAGGGCTCTGCTGCTGCTGCTGCTG	3060
Oy	1921	TAAGCCGCCCTTGTCCTGCCCCATTATTCAAGGGATCTACGGAACTGGCCACTGCC	1980	Db	3001	ATTACTTTRGCGCCTTCAGGAGTACTGGAGGGCTCTGCTGCTGCTGCTGCTG	3060
Db	1921	TAAGCCGCCCTTGTCCTGCCCCATTATTCAAGGGATCTACGGAACTGGCCACTGCC	1980	Qy	3061	ATTCTCGCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3120
Oy	1981	CAGGGACAAAGGTTGGAGTCAAGGCACTAACCGATGATCCAGGAGGACTG	2040	Db	3061	ATTCTCGCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3120
Db	1981	CAGGGACAAAGGTTGGAGTCAAGGCACTAACCGATGATCCAGGAGGACTG	2040	Qy	3121	AGCTTAATACTGGCTTAATTCAGGACCACTGAGCTGAGCTGAGCTGAGCTG	3180
Oy	2040	AGGGTGCCTGGGCAAGGCCATCCATGCGATCACCTCACAGAGCCCTGGTAGCT	2100	Db	3121	AGCTTAATACTGGCTTAATTCAGGACCACTGAGCTGAGCTGAGCTGAGCTG	3180
Db	2040	AGGGTGCCTGGGCAAGGCCATCCATGCGATCACCTCACAGAGCCCTGGTAGCT	2100	Qy	3181	TCCAGCTTACACTGGCTTACCTCATGCAAAACCTTAAGCTAACARARGCTCA	3240
Oy	2101	GACCATGAGGGCAAGGAGTGTGCTCTGGAGACTCTGGGGCTCTGCTGCTG	2160	Db	3181	TCCAGCTTACACTGGCTTACCTCATGCAAAACCTTAAGCTAACARARGCTCA	3240
Db	2101	GACCATGAGGGCAAGGAGTGTGCTCTGGAGACTCTGGGGCTCTGCTGCTG	2160	Qy	3241	GCRTAAVAGGYCTGCGAGAAATGGATCCTCCAGGWGTCGCGAACAGGCGT	3300
Oy	2161	CTCTGTCGGGCAACTGCTCCAGATCTGTCCTACTCTGGGGCTCTGCTGCTG	2220	Db	3241	GCRTAAVAGGYCTGCGAGAAATGGATCCTCCAGGWGTCGCGAACAGGCGT	3300
Db	2161	CTCTGTCGGGCAACTGCTCCAGATCTGTCCTACTCTGGGGCTCTGCTGCTG	2220	Qy	3301	WATCAGTAAATTAGGAAACTCGAACCTTAAATGAGTGGYACNTGA	3360
Oy	2221	GGCAGTCACTAGATACTTTCAGCCACTAAGTTAGTACTGGGGAGCTTATCTT	2280				

QY	5521 ATTCAGCGCTCTTCAACACACAGGAGAAGACTAAATCTAAATCCCC	5580	Db	6601 ACAAACCTCTACTCAGTTCTACTACAACATACTCAAGAACATAATGGGACATGGAAAGG	6660
Db	5521 ATTCAGCGCTCTTCAACACACAGGAGAAGACTAAATCTAAATCCCC	5580	QY	6661 GTCGCCAACTCCCTGGCACCTTGCAAGATCACTAACTCCCTAGGGAGTAGTCCTT	6720
QY	5581 ATGGSCCTCCTTATCATATTTCCTCTKASHTTSTTYACCTSTTCACT	5640	Db	6661 GTCGCCAACTCCCTGGCACCTTGCAAGATCACTAACTCCCTAGGGAGTAGTCCTT	6720
Db	5581 ATGGSCCTCCTTATCATATTTCCTCTKASHTTSTTYACCTSTTCACT	5640	QY	6721 CAAATGAGAGCTTCTAGCTGCTAACGCTGARAGAGGGAACTGTATTGTA	6780
QY	5641 GCACCCCTTCATGCCGCTGATACAGTAGTCCTCCYYACMAGAGTCTATGAGA	5700	Db	6721 CAAATGAGAGCTTCTAGCTGCTAACGCTGARAGAGGGAACTGTATTGTA	6780
Db	5641 GCACCCCTTCATGCCGCTGATACAGTAGTCCTCCYYACMAGAGTCTATGAGA	5700	QY	6781 CGGAAAGATGCTGTTTATGTTAACTCGGAATCGGATGCGCTACTGAGAAGGTTA	6840
QY	5701 ATTCAGCGCTCCGAAATATTGTCGCCCTCTGATAGGGTCTTAAAGGACCCC	5760	Db	6781 CGGAAAGATGCTGTTTATGTTAACTCGGAATCGGATGCGCTACTGAGAAGGTTA	6840
Db	5701 ATTCAGCGCTCCGAAATATTGTCGCCCTCTGATAGGGTCTTAAAGGACCCC	5760	QY	6841 ATTCAGCGCTCCGAAATATTGTCGCCCTCTGATAGGGTCTTAAAGGACCCC	6900
QY	5761 ACCCTACTGCCAACCCATATGCCCTCAACTCTACTTGCCCTCTTGAG	5820	Db	6841 ATTCAGCGCTCCGAAATATTGTCGCCCTCTGATAGGGTCTTAAAGGACCCC	6900
Db	5761 ACCCTACTGCCAACCCATATGCCCTCAACTCTACTTGCCCTCTTGAG	5820	QY	6901 CTTCAGCCTATGCGCTGCTGGATTCACCTGCTCCCTCGGGCTC	6930
QY	5821 CATGCAAATCTATTTGACAGGAAATAGGTTAACTCTAGTGTGCTGGAGCT	5880	Db	6901 CTTCAGCCTATGCGCTGCTGGATTCACCTGCTCCCTCGGGCTC	6930
Db	5821 CATGCAAATCTATTTGACAGGAAATAGGTTAACTCTAGTGTGCTGGAGCT	5880	QY	6961 CAACTCTTGGACCTGATCTTRACCTCTGTTAACCTGCTCTCCAGATC	7020
QY	5881 GGACTCACTCTCTGAGTACTTCACCCAACCTGGTGTGTTGGGGGAGGT	5940	Db	6961 CAACTCTTGGACCTGATCTTRACCTCTGTTAACCTGCTCTCCAGATC	7020
Db	5881 GGACTCACTCTCTGAGTACTTCACCCAACCTGGTGTGTTGGGGGAGGT	5940	QY	7021 GAACTGTTAACTACAAATGGGACCCAGATCAGGCAAGACTGACCTACCGA	7080
QY	5941 CAAGATCAGGAGGAGAAACAGTAAAGGAAATGTTAAAGGAACTCCCGTAC	6000	Db	7021 GAACTGTTAACTACAAATGGGACCCAGATCAGGCAAGACTGACCTACCGA	7080
Db	5941 CAAGATCAGGAGGAGAAACAGTAAAGGAAATGTTAAAGGAACTCCCGTAC	6000	QY	7081 CCCCTGGACCCGCTCTGAGCCACGCTGTTGAGCTAAAGGACCCCTCT	7140
QY	6001 GGCACCTCTTGGCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCGTAC	6060	Db	7081 CCCCTGGACCCGCTCTGAGCCACGCTGTTGAGCTAAAGGACCCCTCT	7140
Db	6001 GGCACCTCTTGGCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCGTAC	6060	QY	7141 GAGGAATCTCACTGCAACCTCTTACGCCCCATTCTACGGAGGAGCTGAGC	7200
QY	6061 CATACTCGCTGTTAACCCATTAACTACCCCTACTGGCTCATGGGTCTCGGC	6120	Db	7141 GAGGAATCTCACTGCAACCTCTTACGCCCCATTCTACGGAGGAGCTGAGC	7200
Db	6061 CATACTCGCTGTTAACCCATTAACTACCCCTACTGGCTCATGGGTCTCGGC	6120	QY	7201 GGTGGTCGACCAACCTCCCAACAGCAGCTTAGGTTCTCTGTTGAGATGGGAGCTGAG	7260
QY	6121 CAAACCTCTACTAAGTGTGATATGCCCTCCCTGAACTTCARGCCATTTGTTCAAC	6180	Db	7201 GGTGGTCGACCAACCTCCCAACAGCAGCTTAGGTTCTCTGTTGAGATGGGAGCTGAG	7260
Db	6121 CAAACCTCTACTAAGTGTGATATGCCCTCCCTGAACTTCARGCCATTTGTTCAAC	6180	QY	7261 GACGAGCTACTGCTGTTCTAGGTTGATAGGATTAACCTCTGAGCTGTTGAGCTGAG	7320
QY	6181 CCTGTAACCTGACATGGAACACTTCAGAGAAATAAACCCACTCGTTGATGAA	6240	Db	7261 GACGAGCTACTGCTGTTCTAGGTTGATAGGATTAACCTCTGAGCTGTTGAGCTGAG	7320
Db	6181 CCTGTAACCTGACATGGAACACTTCAGAGAAATAAACCCACTCGTTGATGAA	6240	QY	7321 ACCACATCCACCTTAAACCGGGGCTTGCACTTAGTGTACACCTGACCAATGAGAG	7380
QY	6241 GGACCTCTTGTCCATSTGAAATAACCATCACCTCACCTACCTGTAATTT	6300	Db	7321 ACCACATCCACCTTAAACCGGGGCTTGCACTTAGTGTACACCTGACCAATGAGAG	7380
Db	6241 GGACCTCTTGTCCATSTGAAATAACCATCACCTCACCTACCTGTAATTT	6300	QY	7381 CTCACTAATGCTTAATGGCAAGACGAGCAGGAGTAAGAAATGCAATCATVTTATG	7440
QY	6301 AGCAACTACTGACACAACTCCATGCACTGGTTACTCTCCACAA	6360	Db	7381 CTCACTAATGCTTAATGGCAAGACGAGCAGGAGTAAGAAATGCAATCATVTTATG	7440
Db	6301 AGCAACTACTGACACAACTCCATGCACTGGTTACTCTCCACAA	6360	QY	7441 MTGAGAGCACAGCAGGAGGAACTGATGGGTTATACCAAGTYYTGGAGCGGCAA	7500
QY	6361 ATAGCTGCTTACCTCAGGAAATTTCCTGTTGACTCTGCTTCAGCTGTTGTT	6420	Db	7441 MTGAGAGCACAGCAGGAGGAACTGATGGGTTATACCAAGTYYTGGAGCGGCAA	7500
Db	6361 ATAGCTGCTTACCTCAGGAAATTTCCTGTTGACTCTGCTTCAGCTGTTGTT	6420	QY	7501 CGGACACCCCTTGGGCCCCCTTGATGGAGCTGTTCTAGCTTATTCAC	7560
QY	6421 AATGGCTCTCAGAACTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT	6480	Db	7501 CGGACACCCCTTGGGCCCCCTTGATGGAGCTGTTCTAGCTTATTCAC	7560
Db	6421 AATGGCTCTCAGAACTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT	6480	QY	7561 TCTATTAATCTGCARCTGCR	7582
QY	6481 ACTGACACAAATTACAGTATGCACTCTAACCCGCCAACAAAGTACCAATT	6540	Db	7561 TCTATTAATCTGCARCTGCR	7582
Db	6481 ACTGACACAAATTACAGTATGCACTCTAACCCGCCAACAAAGTACCAATT	6540	RESULT 4		
QY	6541 CTCCCTTGTATAGGAGGGAGGTGCTGGTCACTGGTACTGGCATGGCGTATC	6600	AX027480		
Db	6541 CTCCCTTGTATAGGAGGGAGGTGCTGGTCACTGGTACTGGCATGGCGTATC	6600	LOCUS	Sequence 30 from Patent FR2788784.	
QY	6601 ACAACCTCTACTCAAACTACTCAAACTAACTGGGACATGGGAAACGG	6660	DEFINITION		
			ACCESSION	AX027480	

VERSION	AX027480.1	GI:	10188444
KEYWORDS			
SOURCE	Homo sapiens		(human)
ORGISM	Bukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	Mammalia; Butheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mallet, F., Voisset, C. and Paranhos, B. G.		
AUTHORS	PATENT: FR 2788784-A 30 28-JUL-2000;		
JOURNAL	BIO MERIEUX (FR)		
FEATURES	Location/Qualifiers		
source	1..7582		
ORIGIN			
Query Match	98.8%; Score 7489.6; DB 6; Length 7582;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches	7582; Conservative		
	Mismatches 0; Indels 0; Gaps 0;		
Qy	1 CAACAATGGGATAAACCCAGGCATTGGACCTGCCAACAGCAGCCCCCTTGGGCC 60		
Db	1 CAACAATCGGATATAACCACSGATTCGAGCTGGCACACAGCAGCCCCCTTGGGCC 60		
Qy	61 CTCCTCTGGTGTGGGGTTCTACGGCTCGAGCTTGTGTCACCTGCA 120		
Db	61 CTCCTCTGGTGTGGGGTTCTACGGCTCGAGCTTGTGTCACCTGCA 120		
Qy	121 CTCTCTGGTGTGGGGTTCTACGGCTCGAGCTTGTGTCACCTGCA 120		
Db	121 CTCTCTGGTGTGGGGTTCTACGGCTCGAGCTTGTGTCACCTGCA 120		
Qy	181 TGTGTGCCAACACGCGAACACTSGCGCTGACTCCATCCTCTGATCTGAGGTCTC 240		
Db	181 TGTGTGCCAACACGCGAACACTSGCGCTGACTCCATCCTCTGATCTGAGGTCTC 240		
Qy	301 CATTGTCCTGCAAGGGCTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240		
Db	301 CATTGTCCTGCAAGGGCTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240		
Qy	241 CGCTGTGCTCTGATCCAGGGARGGGCCATTGGGCTCCAAATGGGTAAAGGCTGC 300		
Db	241 CGCTGTGCTCTGATCCAGGGARGGGCCATTGGGCTCCAAATGGGTAAAGGCTGC 300		
Qy	361 GCGTCCATGTTCTCTGTGACCCACGGCTCTAATAKACTTAACTTACCTACCA 420		
Db	361 GCGTCCATGTTCTCTGTGACCCACGGCTCTAATAKACTTAACTTACCTACCA 420		
Qy	421 TGCCCAAAGATTCATTCTGGATCCCTGGGCTGAGGAACTCCAGGTAGAGATA 480		
Db	421 TGCCCAAAGATTCATTCTGGGCTGAGGAACTCCAGGTAGAGATA 480		
Qy	481 GARGCTTGCCACCATCTGGAAAGGGCCCTGCTACCTCTGGAGTGTCTTACCATC 540		
Db	481 GARGCTTGCCACCATCTGGAAAGGGCCCTGCTACCTCTGGAGTGTCTTACCATC 540		
Qy	541 TGGGAGSTCTGTGAGCAAGGACCCGGTRACATTGGACACANSRACGACATCC 600		
Db	541 TGGGAGSTCTGTGAGCAAGGACCCGGTRACATTGGACACANSRACGACATCC 600		
Qy	601 MAAGTGTGGAAACGTTCCCGGAGCAAAAGGCCCTAACAGCGTATTCTGGARAT 660		
Db	601 MAAGTGTGGAAACGTTCCCGGAGCAAAAGGCCCTAACAGCGTATTCTGGARAT 660		
Qy	661 TGGGACATTGGACCCCTGAGCACTAAGAAGAACCTTAATCTCTGGACGCC 720		
Db	661 TGGGACATTGGACCCCTGAGCACTAAGAAGAACCTTAATCTCTGGACGCC 720		
Qy	721 GCTGGACTCTGAGGAGTAAATTAAACACATCTTCAGCTAGCYCTTGT 780		
Db	721 GCTGGACTCTGAGGAGTAAATTAAACACATCTTCAGCTAGCYCTTGT 780		
Qy	1861 TAGGGCTTGCTCCAGGGCGGTCTACAGGACACTTAAAGATTTGTCAGTAGAAG 1920		

Db	1861	 TAGGGCTTGCTTCAGTGGCTCAGAACACTTAAAGATGTCAGTAGAG	QY	1920	 1921 TAGCCGCCCTCTGCCCTTAATTCAGGGATCACTGGAGGCCACTGCC	Db	1980	 1981 TRAGCCGCCCTCTGCCCTTAATTCAGGGATCACTGGAGGCCACTGCC	QY	1980	 1981 CAGGGACAAAGGTCTTGTAGTCAGAACCACTAACCGATGATCCAGGGACTG	Db	2040	 1981 CAGGGACAAAGGTCTTGTAGTCAGAACCACTAACCGATGATCCAGGGACTG	QY	2040	 2041 AGGGTGCCTGGCAAGGCCATCCATGCCATCACCC	Db	2100	 2041 AGGGTGCCTGGCAAGGCCATCCATGCCATCACCC	QY	2100	 2101 GACCATGGGGCAGGAAGGGTCTCTGGACACTGGTCTCTAGTTACT	Db	2160	 2101 GACCATGGGGCAGGAAGGGTCTCTGGACACTGGTCTCTAGTTACT	QY	2160	 2161 CTTCTGTCGGGAACTGCTTCAGATGTGACTACTCTGAGGCTCTAGG	Db	2220	 2161 CTTCTGTCGGGAACTGCTTCAGATGTGACTACTCTGAGGCTCTAGG	QY	2220	 2211 GGAGTCACTAGATACTTTCAGGCACTAAGCTGCTGAGCTGACTACTCTGAGG	Db	2280	 2211 GGAGTCACTAGATACTTTCAGGCACTAAGCTGCTGAGCTGACTACTCTGAGG	QY	2280	 2221 GGCAGTCACTAGATACTTTCAGGCACTAAGCTGCTGAGGCTCTAGCT	Db	3000	 2221 GGCAGTCACTAGATACTTTCAGGCACTAAGCTGCTGAGGCTCTAGCT	QY	3000	 3001 GCTAAVAGGYCTGCGGAATAGGGATCTCCAGGTGGCTGCAAATAGCC	Db	3060	 3001 GCTAAVAGGYCTGCGGAATAGGGATCTCCAGGTGGCTGCAAATAGCC	QY	3060	 3061 ATTCTCTGCGYACCTGTGGCTACAWGGTTCCAACCSARAGCTCARCTCTGCTCAGC	Db	3120	 3061 ATTCTCTGCGYACCTGTGGCTACAWGGTTCCAACCSARAGCTCARCTCTGCTCAGC	QY	3120	 3121 AGTTAAATACTTAGGCTAARRATTACCAAGGACACARGCCCTGAGTGA	Db	3180	 3121 AGTTAAATACTTAGGCTAARRATTACCAAGGACACARGCCCTGAGTGA	QY	3180	 3181 TCCAGCTTAACTGGCTATCTCTGAGCTGAGGCTCTCTAGTGTAGT	Db	3240	 3181 TCCAGCTTAACTGGCTATCTCTGAGCTGAGGCTCTCTAGTGTAGT	QY	3240	 3241 GCTAAVAGGYCTGCGGAATAGGGATCTCCAGGTGGCTGCAAATAGCC	Db	3300	 3241 GCTAAVAGGYCTGCGGAATAGGGATCTCCAGGTGGCTGCAAATAGCC	QY	3300	 3301 WATACASTPAATTAGGAAACTCGAGAACGCAAATACCCATTARTAAGTGGAYAMCIGA	Db	3360	 3301 WATACASTPAATTAGGAAACTCGAGAACGCAAATACCCATTARTAAGTGGAYAMCIGA	QY	3360	 3361 GMYRAAGTGGCTTCCAGGCCCTAAAGAAGGCTTAACCCRAYGCCCCAGT	Db	3420	 3361 GMYRAAGTGGCTTCCAGGCCCTAAAGAAGGCTTAACCCRAYGCCCCAGT	QY	3420	 3421 TGCCAAAGGGCAAGCTTSTIAYRTCAAGAAAAAACGRAYASCTCTGGA	Db	3480	 3421 TGCCAAAGGGCAAGCTTSTIAYRTCAAGAAAAAACGRAYASCTCTGGA	QY	3480	 3481 GTCTTACAGTCAGTCCAGGGAYGAGCTGCAACCTTGCCRACYRGTAA	Db	3540	 3481 GTCTTACAGTCAGTCCAGGGAYGAGCTGCAACCTTGCCRACYRGTAA	QY	3540	 3541 GATGTTAGGCAAGGGTGRCYTCATGTTAGGGTAGTGGCTGGAGTAGT	Db	3600	 3541 GATGTTAGGCAAGGGTGRCYTCATGTTAGGGTAGTGGCTGGAGTAGT	QY	3600	 3601 GTATCTGAGGAGTTAAATACAGGAGATCTACTGTTGCACTCTCATGAK	Db	3660	 3601 GTATCTGAGGAGTTAAATACAGGAGATCTACTGTTGCACTCTCATGAK	QY	3660	 3660 3661 GTGAAVRGCACTACCTCTAAAGGAGACTGTTGCTGAGCACAGCTTACTAA	Db	3720	 3661 GTGAAVRGCACTACCTCTAAAGGAGACTGTTGCTGAGCACAGCTTACTAA	QY	3720	 3721 TRTCAGCTCTATTACTTCAGGCGCACTGCTGCRACACTGCACTGTTACTAA	Db	3780	 3721 TRTCAGCTCTATTACTTCAGGCGCACTGCTGCRACACTGCACTGTTACTAA	QY	3780	 3780 3781 CCAGYONCATTTCTCCAGCAATGAGAATRAYAATAGTCACAATT	Db	3840	 3781 CCAGYONCATTTCTCCAGCAATGAGAATRAYAATAGTCACAATT	QY	3840	 3781 CCAGYONCATTTCTCCAGCAATGAGAATRAYAATAGTCACAATT	Db	3840	 3781 CCAGYONCATTTCTCCAGCAATGAGAATRAYAATAGTCACAATT	QY	3840	 3841 TCTCAACCTATGCCACTGGAGGACCTTGTAGGTTCTGCTGCACTG	Db	3900	 3841 TCTCAACCTATGCCACTGGAGGACCTTGTAGGTTCTGCTGCACTG	QY	3900	 3901 CAACGTGTTACTGAGGAGTCTTGTAGAAAAGGACTCGAACAGYGGGTATGC	Db	3960	 3901 CAACGTGTTACTGAGGAGTCTTGTAGAAAAGGACTCGAACAGYGGGTATGC	QY	3960	 3901 CAACGTGTTACTGAGGAGTCTTGTAGAAAAGGACTCGAACAGYGGGTATGC	Db	3960	 3961 AGTGTCACTGATAATGGAAATAYTGTAGAATTCCTCTACTCGGAACTAGTGTCA	Db	4020	 3961 AGTGTCACTGATAATGGAAATAYTGTAGAATTCCTCTACTCGGAACTAGTGTCA	QY	4020	 4021 GCTGGAGAACTATAGCCYTCAVKGCCACTAGTGAATTAGGAGAGRAAAAGGYAA	Db	4080	 4021 GCTGGAGAACTATAGCCYTCAVKGCCACTAGTGAATTAGGAGAGRAAAAGGYAA	QY	4080
----	------	--	----	------	---	----	------	--	----	------	--	----	------	--	----	------	---	----	------	---	----	------	--	----	------	--	----	------	---	----	------	---	----	------	---	----	------	---	----	------	--	----	------	--	----	------	--	----	------	--	----	------	---	----	------	---	----	------	---	----	------	---	----	------	--	----	------	--	----	------	--	----	------	--	----	------	---	----	------	---	----	------	--	----	------	--	----	------	---	----	------	---	----	------	--	----	------	--	----	------	--	----	------	--	----	------	---	----	------	---	----	------	--	----	------	---	----	------	--	----	------	--	----	------	--	----	------	---	----	------	---	----	------	---	----	------	--	----	------	--	----	------	--	----	------	--	----	------	--	----	------	---	----	------	---	----	------	---	----	------	---	----	------

Oy	4081	TATTTATACAGACTCTRARTATGTYACCTAGTCCTCATGCCCATGMGCATATGSR	4140
Db	4081	TATATATACAGACTCTRARTATGTYACCTAGTCCTCATGCCCATGMGCATATGSR	4140
Oy	4141	AGAAGGGATTCTTAACTCYGGAGAACCTATCAGGAAGCCATTAGGAACTTACGAG	4200
Db	4141	AGAAGGGATTCTTAACTCYGGAGAACCTATCAGGAAGCCATTAGGAACTTACGAG	4200
Oy	4201	ATTATAYGGCNGTACAGAACCTARAGAGGTGAGCTTACAGTGCTGTCATCA	4260
Db	4201	ATTATAYGGCNGTACAGAACCTARAGAGGTGAGCTTACAGTGCTGTCATCA	4260
Oy	4261	NAAGGAGAAGRAAGGAAATASAGRGAYTGCACAAKATATGAGCMMAAGGC	4320
Db	4261	NAAGGAGAAGRAAGGAAATASAGRGAYTGCACAAKATATGAGCMMAAGGC	4320
Oy	4321	TGCAAGGGAGACCTCCATTAGAATGTTATAACTCCCTTGTAGTAGGTTATCC	4380
Db	4321	TGCAAGGGAGACCTCCATTAGAATGTTATAACTCCCTTGTAGTAGGTTATCC	4380
Oy	4381	CTTCCCAGGAAACCAAAGCCCGTACAGTCAAGCTCAAGG	4440
Db	4381	CTTCCCAGGAAACCAAAGCCCGTACAGTCAAGCTCAAGG	4440
Oy	4441	CACTTTCCTCCCTCGGGACGGTGTAGCCACTGAGAAGGAAATCTTGCCTGCAC	4500
Db	4441	CACTTTCCTCCCTCGGGACGGTGTAGCCACTGAGAAGGAAATCTTGCCTGCAC	4500
Oy	4501	TATCCAACTGAAATTACTTAAACCTCTCATCAAACCTTCTAGCATCCAGCAC	4560
Db	4501	TATCCAACTGAAATTACTTAAACCTCTCATCAAACCTTCTAGCATCCAGCAC	4560
Oy	4561	AKTCAGGGCTCTGTCAGTGTGCCRARAATAATCCCCGCTTATGCCAAGCTCC	4680
Db	4561	AKTCAGGGCTCTGTCAGTGTGCCRARAATAATCCCCGCTTATGCCAAGCTCC	4680
Oy	4621	CCATCARATGGCCAATCATATTATTTACTGGACCAGGCCTTCAACATCAGCARAT	4620
Db	4621	CCATCARATGGCCAATCATATTATTTACTGGACCAGGCCTTCAACATCAGCARAT	4620
Oy	4681	AGGAAACAAARACAGGCCATTACCTGTRARARACTGCGCACTGTTACCCACAG	4740
Db	4681	AGGAAACAAARACAGGCCATTACCTGTRARARACTGCGCACTGTTACCCACAG	4740
Oy	4741	CCCAACCTCAGGTTCACTACTGCTGGTARATACTTACGGGTTGGCA	4800
Db	4741	CCCAACCTCAGGTTCACTACTGCTGGTARATACTTACGGGTTGGCA	4800
Oy	4801	RAGGCCTTCCCTGAGGACAGAACGGCTTACGGGTTAGGTTATGAGCTTACGGGTTGGCA	4860
Db	4801	RAGGCCTTCCCTGAGGACAGAACGGCTTACGGGTTAGGTTATGAGCTTACGGGTTGGCA	4860
Oy	4861	ATTAATCCAGATTGGACTTCCCGAGGCTTCAGACTGACATASCCCTGTTCCAG	4920
Db	4861	ATTAATCCAGATTGGACTTCCCGAGGCTTCAGACTGACATASCCCTGTTCCAG	4920
Oy	4921	GCCACAGTAACCCAGGGATTCAGGGTTAGGTTAGATATCACTACAGCCTGCC	4980
Db	4921	GCCACAGTAACCCAGGGATTCAGGGTTAGGTTAGATATCACTACAGCCTGCC	4980
Oy	4981	TGAGGCCACAGTCTCAGGAAGCTCATGGCTGTCGTGCTCATAGCTTAAA	5040
Db	4981	TGAGGCCACAGTCTCAGGAAGCTCATGGCTGTCGTGCTCATAGCTTAAA	5040
Oy	5041	AAAGAGAACCCAGGAACCCACTCATGGCTGTCGTGCTCATAGCTTAAA	5100
Db	5041	AAAGAGAACCCAGGAACCCACTCATGGCTGTCGTGCTCATAGCTTAAA	5100
Oy	5101	AGATCTGCAACTTCCCAAAGCAGGACTAGCCATACGAAATCTGTGAGG	5160
Db	5101	AGATCTGCAACTTCCCAAAGCAGGACTAGCCATACGAAATCTGTGAGG	5160
Oy	5161	CCCTCATACACCATGACCTTGTGCTGACAGACGCCACTTAGTGCAGACATCA	5220
Db	5161	CCCTCATACACCATGACCTTGTGCTGACAGACGCCACTTAGTGCAGACATCA	5220
Oy	5221	CCTCCTTAGCCAAATATCACAGTCTTAACACATCAAGAACCTATCCCTGAGA	5280
Db	5221	CCTCCTTAGCCAAATATCACAGTCTTAACACATCAAGAACCTATCCCTGAGA	5280
Oy	5281	AGGAAAGAACATATCCACCGWGTGACATGGTATAGTCAGTCCCTCYCTTAAT	5340
Db	5281	AGGAAAGAACATATCCACCGWGTGACATGGTATAGTCAGTCCCTCYCTTAAT	5340
Oy	5341	CCCATCCCTAACTACATCCTGGAGAACCTTACCCAGTCTTATGCAAGTCCCTCYCTTAAT	5340
Db	5341	CCCATCCCTAACTACATCCTGGAGAACCTTACCCAGTCTTATGCAAGTCCCTCYCTTAAT	5340
Oy	5401	CGTTAAGTGTGCTGGAGTGTGCTGATCATCACACTTGTCATCTGGATAC	5460
Db	5401	CGTTAAGTGTGCTGGAGTGTGCTGATCATCACACTTGTCATCTGGATAC	5460
Oy	5440	CGTTAAAGTGTGCTGGAGTGTGCTGATCATCACACTTGTCATCTGGATAC	5460
Db	5440	CGTTAAAGTGTGCTGGAGTGTGCTGATCATCACACTTGTCATCTGGATAC	5460
Oy	5461	TGCAAAAGGACCTGAAATCCAGGAGCACGCCACTGAGCTTACCTAGCTAGG	5520
Db	5461	TGCAAAAGGACCTGAAATCCAGGAGCACGCCACTGAGCTTACCTAGCTAGG	5520
Oy	5521	ATTTGCGCTGTCTTCAACACAACTGAGGAAATCTTGTGATCTTCTACT	5580
Db	5521	ATTTGCGCTGTCTTCAACACAACTGAGGAAATCTTGTGATCTTCTACT	5580
Oy	5581	AGGSCCTCCCTATCATATTTCCTKTAATGTTTTCCTKTAATGTTTTCCTTCTACT	5640
Db	5581	AGGSCCTCCCTATCATATTTCCTKTAATGTTTTCCTKTAATGTTTTCCTTCTACT	5640
Oy	5641	GAACCCCTCTCCATGCCGCTGTTGACCTAGCTTACATAAATCC	5700
Db	5641	GAACCCCTCTCCATGCCGCTGTTGACCTAGCTTACATAAATCC	5700
Oy	5701	ATCCAGGGTCCGGAAATATCTGTCCTCATGTTTCTCTKTAATGTTTTCCTTCTACT	5760
Db	5701	ATCCAGGGTCCGGAAATATCTGTCCTCATGTTTCTCTKTAATGTTTTCCTTCTACT	5760
Oy	5760	ATCCAGGGTCCGGAAATATCTGTCCTCATGTTTCTCTKTAATGTTTTCCTTCTACT	5760
Db	5760	ATCCAGGGTCCGGAAATATCTGTCCTCATGTTTCTCTKTAATGTTTTCCTTCTACT	5760
Oy	5761	ACCTTCACTGCCACACCCATTGCCCCATGCTGCACTGCTATCATCTGCGCACTTGTGATG	5820
Db	5761	ACCTTCACTGCCACACCCATTGCCCCATGCTGCACTGCTATCATCTGCGCACTTGTGATG	5820
Oy	5821	CATGCAAACTCATCATTTGTCGAGGAAATGATTACCTGCTGCTGCTGCTGCTGCTG	5880
Db	5821	CATGCAAACTCATCATTTGTCGAGGAAATGATTACCTGCTGCTGCTGCTGCTG	5880
Oy	5881	GGACTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5940
Db	5881	GGACTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5940
Oy	5941	CGAGTCACTGAGGAAACATGTAAGAGGAAATCTCCCAACTGCCGGTACAT	6000
Db	5941	CGAGTCACTGAGGAAACATGTAAGAGGAAATCTCCCAACTGCCGGTACAT	6000
Oy	6001	GGACTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6120
Db	6001	GGACTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6120
Oy	6121	CGAACCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6180
Db	6121	CGAACCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6180
Oy	6180	CGAACCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6180
Db	6180	CGAACCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6180
Oy	6181	CTGTACTGCAACTGAGGAAACCTGAGGAAATACCACTCCGTTAGA	6240
Db	6181	CTGTACTGCAACTGAGGAAACCTGAGGAAATACCACTCCGTTAGA	6240

Db 6241 GGACTCTGTTCAAATGGAAATAACACGGGCTGCACCTAGTCACCTGCCAAATT 6300
 Qy 6301 AGCAATCTACATAACACACAACTCCAAATSGATCAGGTGSGTAACTCTCCACCAA 6360
 Db 6301 AGCAATCTACATAACACACAACTCCAAATSGATCAGGTGSGTAACTCTCCACCAA 6360
 Qy 6361 ATAGTCGCTTACCTCAGGAATTATTGTTGCTGTTGACTCAGGTTG 6420
 Db 6361 ATAGTCGCTTACCTCAGGAATTATTGTTGCTGTTGACTCAGGTTG 6420
 Qy 6421 AATGGCTTCAAGATCTAGTGTTGCTATCPAGCCGCHACAAAGAGTACCAATT 6480
 Db 6421 AATGGCTTCAAGATCTAGTGTTGCTATCPAGCCGCHACAAAGAGTACCAATT 6480
 Qy 6481 ACTGAAACAGATTATCAGTTAGTGTCAATCTAGCCCCYATGRCCATCTAC 6540
 Db 6481 ACTGAAACAGATTATCAGTTAGTGTCAATCTAGCCCCYATGRCCATCTAC 6540
 Qy 6541 CTCTCTTGTATAGGAGGAGGTAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 6600
 Db 6541 CTCTCTTGTATAGGAGGAGGTAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 6600
 Qy 6601 ACRAACTCTACTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 6660
 Db 6601 ACRAACTCTACTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 6660
 Qy 6661 GTRGCCGACTCCCTGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 6720
 Db 6661 GTRGCCGACTCCCTGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 6720
 Qy 6721 CAAATCGAGGCTTAGACTCTAACCGTGTGARAGGGGGACCTGTTATTGTA 6780
 Db 6721 CAAATCGAGGCTTAGACTCTAACCGTGTGARAGGGGGACCTGTTATTGTA 6780
 Qy 6781 GGGGAGAGATGCTGTTATTGTTATCCTCGGAATGTCAGTGGAGATRAGAA 6840
 Db 6781 GGGGAGAGATGCTGTTATTGTTATCCTCGGAATGTCAGTGGAGATRAGAA 6840
 Qy 6841 ATCCSAGATCGATAACAGCTAKAGCAGGAGCTCGAACAATGGACCTGGGCCTC 6900
 Db 6841 ATCCSAGATCGATAACAGCTAKAGCAGGAGCTCGAACAATGGACCTGGGCCTC 6900
 Qy 6901 CTCAGGCRATGGATCCCCTCGATTCCTCTAGGACCTCTAGGACCTPATATG 6960
 Db 6901 CTCAGGCRATGGATCCCCTCGATTCCTCTAGGACCTCTAGGACCTPATATG 6960
 Qy 6961 CTACTCTCTTGGACCCCTGATCTTACCTCTCTCTCTCTCTCTCTCTCTCTCT 7020
 Db 6961 CTACTCTCTTGGACCCCTGATCTTACCTCTCTCTCTCTCTCTCTCTCTCTCT 7020
 Qy 7021 GAACTGTTAACTACAAATGGACGCCAGATGCGTCCAGACTAGATCTACCGAGA 7080
 Db 7021 GAACTGTTAACTACAAATGGACGCCAGATGCGTCCAGACTAGATCTACCGAGA 7080
 Qy 7081 CCCCTGGACCCGCTGTYACCCACGATCTGATGTAATGACATCAAGGCCACCTCT 7140
 Db 7081 CCCCTGGACCCGCTGTYACCCACGATCTGATGTAATGACATCAAGGCCACCTCT 7140
 Qy 7141 GAGGAATCTCAGTCAGTCACCTACAGCCAAATTAGCAGGAAACAGCTAGAC 7200
 Db 7141 GAGGAATCTCAGTCAGTCACCTACAGTCAGGAAACAGCTAGAC 7200
 Qy 7201 GGTGSGTGCSCAACCTCCACACGACTTAGTTTCTGTTGAGTGGGACTGAGA 7260
 Db 7201 GGTGSGTGCSCAACCTCCACACGACTTAGTTTCTGTTGAGTGGGACTGAGA 7260
 Qy 7261 GACAGGACTGCTGATTCCTAGGCTGATAGATCTGCTGAGTGGAGGAGG 7320
 Db 7261 GACAGGACTGCTGATTCCTAGGCTGATAGATCTGCTGAGTGGAGGAGG 7320
 Qy 7321 ACCACATCCACCTTAACACGGGCTGCACCTAGTCACCTGCCAAATT 7380

Db 7321 ACCACATCCACCTTAACACGGGCTGCACCTAGTCACCTGCCAAATT 7380
 Qy 7381 CTCACTPAACTCTAATAGCCAAAGCACAGGAGTAAAGAATAGCCAACTATYTATG 7440
 Db 7381 CTCACTPAACTCTAATAGCCAAAGCACAGGAGTAAAGAATAGCCAACTATYTATG 7440
 Qy 7441 MTGAGACACAGCAGGGGAATAGTCGGGATAAACCAAGTTTCAGCCGCAA 7500
 Db 7441 MTGAGACACAGCAGGGGAATAGTCGGGATAAACCAAGTTTCAGCCGCAA 7500
 Qy 7501 CGCACACCCCTTGGTCCCTCCCTTGTATGGGGCTCTGTTCATGTTAAC 7560
 Db 7501 CGCACACCCCTTGGTCCCTCCCTTGTATGGGGCTCTGTTCATGTTAAC 7560
 Qy 7561 TCTATTAATCTGCACTGCR 7582
 Db 7561 TCTATTAATCTGCACTGCR 7582

RESULT 5

AY101582

DEFINITION AY101582 10222 bp DNA linear PRI 11-FEB-2004

LOCUS Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVWE1 locus, allele A, complete sequence.

ACCESSION AY101582

VERSION AY101582.1 GI:37544399

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 10222)

AUTHORS Mallet, F., Bonton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnau, B., Lacombe, G., Duret, L. and Mandrand, B.

TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)

PUBLISHED 14/05/2002

REFERENCE 2 (bases 1 to 10222)

AUTHORS Mallet, F., Bonton, O. and Oriol, G.

TITLE Direct Submission

JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France

FEATURES

source

1. .10222 /organism="Homo sapiens"

/mol_type="genomic DNA"
 /isolate="21"
 /isolation_source="BMC of individual 21"
 /db_xref="Taxon:9605"
 /chromosome="7"
 /map="7q21-q22"
 /sex="female"
 /cell_type="PBMC"
 /note="isolated by PCR; ERVWE1 locus, allele A
 endogenous virus: HERV-W"
 1. .700
 /note="putative mRNA transcript 1; gag-pol pseudogene"
 /note="putative mRNA transcript 1; gag-pol pseudogene"
 /join="248..811,2887..9813)
 /join="248..811,758..9813)
 /product="envelope glycoprotein"
 /note="putative mRNA transcript 2"
 /join="248..811,9248..9813)
 /note="putative mRNA transcript 3; no coding region"
 7812..9428
 /note="synonym"
 /codon_start=1
 /product="envelope glycoprotein"
 /protein_id="PAM68161_1"
 /ab_xref="GI:37544400"
 /translation="MPLVHFLFTVLPSPFLAPPCCRMTSSSPYQEFARMOR
 GNIDAPSYRSLSKGTPTAHTHMPRNCYHSATLCMHANTHYWTGMINPSCPGGLV

LTR		TVCWHTFTQTMMSDGGVVDDOAREKHEVKEVISOLTRVHGTSRSPYKGUDLSKUHETRT									
ORIGIN	Query	Match	Score	DB	Length	10222;	Best Local Similarity	95.7%	Pred.	No.	Mismatches
	Matches	6681;	Conservative	95.7%;	pred.	50;	Indels	53;	Gaps	27;	
	Qy	606	GATGGAAACGTTCCCGCAAGACAACAGGCCCTTACGACTATTCTGARATTGGA	665	Db	3776	-GGGAGGGAGAAAATCTACTGCTTCAGAGACTAAGGGCATGGAGCAGTGAGAAC	3834			
	Db	2886	GATGGAAACGTTCCCGCAAGACAACAGGCCCTTACGACTATTCTGAGATGGGA	2945	Qy	1566	GTCCTCTCTGTCACITGACTCTCTGAAAGGCCACTAACTTAAGGTAATTATCA	1625			
	Qy	666	MCAATTGACCTTCAACACTAAGAAGAACGACTTATATCTTGCACTGCCCTG	725	Db	2946	RGCTCLFGLRECCYVNOGLVTEKTEIDRORAEELNTGWGLSOMMPRL	9487.	RSDVNIDKPPPERIBISAQPLPNSAGSS"	10222	
	Db	2946	CCAATTGACCTTCAACACTAAGAAGAACGACTTATATCTTGCACTGCCCTG	3005	Qy	1745	FLGFLAKAIIILFGFCINFLVNVSSLRALKQMEPKQSKKLYRPLDRASP				
	Qy	726	GCACCTCTGAGGGAAACTATAATTACACATCTACACCTAGCYCTTGTGAGA	785	Db	3835	GTCCTCTCTGTCACITGACTCTCTGAGGCCACTAACTTAAGGTAATTATCA	3894			
	Db	3006	GCACCTCTGAGGGAACTATAATTACACATCTACACCTAGCYCTTGTGAGA	3065	Qy	1626	CTCAGTCAGCTGAGCAGATTAGAAAACCTCAAGTCTCCGTTGGCGGAGAA	1684			
	Qy	786	AAGGAAATGGAGTAGTGAAGTGCATAAGTACAACTTCTTCAATTAGAGAAC	845	Db	3895	CTCAGTCAGCTGAGCAGATTAGAAAACCTCAAGTCTCCGTTGGCGGAGAA	3954			
	Db	3066	AAGGAAATGGAGTAGTGAAGTGCATAAGTACAACTTCTTCAATTAGAGAAC	3125	Qy	1745	ACTAGAACCTTGGAGCTCTGAGATAAAAGGAAAGTGGCAATATGATGAC	1804			
	Qy	846	AATTAGTGTAAAAGTGTGATTATGGCTCTAGGAGGCTTCAAGAGACACTC	905	Db	4015	AGCGGAAACGAGAACAGGAACTAAAGGAAAGTGGCAATATGATGAC	4074			
	Db	3126	AATTAGTGTAAAAGTGTGATTATGGCTCTAGGAGGCTTCAAGAGACACTC	3185	Qy	1805	CAAGTGAACCTTGGAGCTCTGAGATAAAAGGAAAGTGGCAATATGATGAC	1864			
	Qy	906	CCAGATTCGGCACTCTTCCCACMYTAATAGGACCCCTTCACCCAAATGGCAA	965	Db	4075	CTAGGACTTTGGAGCTCTGAGATAAAAGGAAAGTGGCAATATGATGAC	4134			
	Db	3186	CCAGCTCCCGACTCTTCCCACACTAATAGGACCCCTTCACCCAAATGGCAA	3245	Qy	1865	ACTAGAACCTTGGAGCTCTGAGATAAAAGGAAAGTGGCAATATGATGAC	1744			
	Qy	966	AAGGAGATAGAACAAAGGTTAACAGTGACCAAGAGTGCACATTCCCAATATGA	1025	Db	4135	ACTTGTAAACCTTGTGACTCTGGCTACCTGGTCTTAAAGATGTCAGAAGGAGC	4194			
	Db	3246	AAGGAGATAGAACAAAGGGTAACAGTGACCAATTTCCCAATATGA	3305	Qy	1925	CGCCCTTCCTCATGCCCTTATTCAGGGAACTACTGGAGGCCACTCCCCAG	1984			
	Qy	1026	CCCTCCCAAGCAGCTGGAGGAGAGAAATTGGCCAGCCAGGAGTCACTGGCT	1085	Db	4195	CGCCCTTCCTCATGCCCTTATTCAGGGAACTACTGGAGGCCACTCCCCAG	4254			
	Db	3306	CCCTT-CCAAAGCAGTGTTGGAGGAGAGAAATTGGCCAGCCAGGAGTCACTGGCT	3364	Qy	1985	GGACAAGGTGTTCAGTCAGAGGAACTAACAGATGATGATGAC	2044			
	Qy	1086	TCTCCAGACTTAAGCAATTAAACAGACTTGTGAAATTCTCGATAAYCTGTATG	1145	Db	4255	GGACAGGGCTCTGAGGAACTAACAGGAACTAACAGATGATGATGAC	4314			
	Db	3365	TCTCCAGACTTAAGCAATTAAACAGACTTGTGAAATTCTCGATAACCTGTATG	3424	Qy	2045	TGCTCGGGAGCSCCATCCATCCATCACCTCACAGAGCCCTGGTAGCTGAC	2104			
	Qy	1146	CTATATGTTGTTAACAGGGTTAGGACAACTTCTGATCTGACATGGAGAGATATA	1205	Db	4315	TGCTCGGGAGCSCCATCCATCACCTCACAGAGCCCTGGTAGCTGAC	4374			
	Db	3425	CTATATGTTGTTAACAGGGTTAGGACAACTTCTGATCTGACATGGAGAGATATA	3483	Qy	2105	ATGAGGGCCAGGAAGGTGTCCTGGACACTGGGGGCTTCTAGCTTACTCTC	2164			
	Qy	1206	TGTCACTGCTAACTGACACTAACCCAAATGAGGAAAGGCCACACTACCTG	1265	Db	4375	ATGAGGGCCAGGAAGGTGTCCTGGACACTGGGGGCTTCTAGCTTACTCTC	4433			
	Db	3484	TGTCACTGCTAACTGACACTAACCCAAATGAGGAAAGGCCACACTACCTG	3453	Qy	2165	TGTCGGAGAACACTGGCTCGACATGTCGACTTCTGAGGGGCTNTAGACGGGCA	2224			
	Qy	1266	GAGAGTTGGCATCTGGPATCTGAGTCAGGTCAATGGATGAGAAGTGCAC	1325	Db	4434	TGTCGGAGAACACTGGCTCGACATGTCGACTTCTGAGGGGCTTCTAGGGCA	4491			
	Db	3544	GAGAGTTGGCATCTGGPATCTGAGTCAGGTCAATGGATGAGAAGTGCAC	3600	Qy	2225	GTCACTAGATACTTTCAGCCCTAAGTTGACTGGAGCTTATCTTTCAC	2284			
	Qy	1326	AAAGAGAAATGATTCGCCACAGGCCACGCGAGTCAGCTASACCCCTGAGGGAC	1385	Db	4492	GTCACTAGATACTTCTCCAGCCACTAAGTTGACTGGAGCTTATCTTTCAC	4549			
	Db	3601	AAAGAGAAATGATTCGCCACAGGCCACGCGAGTCAGCTAACCTGAGGAG	3657	Qy	2285	ATGTTTCTAAATAGCTGAGAACGCCACACTACCTGTTAGGGAGACATCTAGCA	2344			
	Qy	1386	ACGAGAAATGATTCGCCACAGGCCACGCGAGTCAGCTAACCTGAGGAG	1445	Db	4550	ATGCTTTCATATGCTGAGAACGCCACACTACCTGTTAGGGAGACATCTAGCA	4609			
	Db	3658	ACGAGAAATGATTCGCCACAGGCCACGCGAGTCAGCTAACCTGAGGAG	3716	Qy	2345	AGCAGGGCCATTAATACACTGACATAGGAAAGGCCACCTCTTCTGAGCAAGGA	2404			
	Qy	1446	GGACTTAACTGAGAAACTGATGAACTTCAATGAGTCCACCATACACA	1505	Db	4610	AGCAGGGCCATTAATACACTGACATAGGAAAGGCCACCTCTTCTGAGCAAGGA	4668			
	Db	3544	GGACTTAACTGAGAAACTGATGAACTTCAATGAGTCCACCATACACA	3600	Qy	2405	CTGAGGAAGGAAATATCTGAGCTGGCCACGAGAACGACANTATGGAGGCCAA	2464			
	Qy	1326	AAAGAGAAATGATTCGCCACAGGCCACGCGAGTCAGCTASACCCCTGAGGGAC	1385	Db	4669	CTGAGGAAGGAAATATCTGAGCTGGCCACGAGAACGACANTATGGAGGCCAA	4727			
	Db	3601	AAAGAGAAATGATTCGCCACAGGCCACGCGAGTCAGCTAACCTGAGGAG	3657	Qy	2465	GATGCCGCTCTGTCAGTTAACTGAGAACGACANTATGGAGGCCAA	2524			
	Qy	1386	ACGAGAAATGATTCGCCACAGGCCACGCGAGTCAGCTAACCTGAGGAG	1445	Db	4728	GTACCCCTGAGGCCACAGGCCACGAGAACGACANTATGGAGGCCAA	4787			
	Db	3658	ACGAGAAATGATTCGCCACAGGCCACGCGAGTCAGCTAACCTGAGGAG	3716	Qy	2525	GTACCCCTGAGGCCACAGGCCACGAGAACGACANTATGGAGGCCAA	2584			
	Qy	1446	GGACTTAACTGAGAAACTGATGAACTTCAATGAGTCCACCATACACA	1505	Db	4788	GTACCCCTGAGGCCACAGGCCACGAGAACGACANTATGGAGGCCAA	4847			

QY	2585	CCAGAGCTTAGTAAACCACTGCAACTCCTGAGTAATCCCTAGTGGAGGAGG	2644	QY	3665	ATRGCACTACTGCAAGAAGACTGTGAGCTGAGACACYSITTAATTC	3724
Db	4848	CCANGGCCTAGTAAACCACTGCAACTCCTGAGTAATCCCTGAGACTCC	4900	Db	5897	ATGGCATACTACTGCAAGAAGACTGTGAGCTGAGACACYSITTAATTC	5956
QY	2645	CACAGAACCCAGTGACAGTGGAGGGTTAGTSGAGATCTAGGATATCAA	2704	QY	3725	AGCTCTTAACTCTGAGGGCGAGTGTGCAACTCTAACCC	3784
Db	4901	TACAGAACCCACAGACAGTGGAGTGGAGATCTAGGATATCAA	4958	Db	5957	AGCTCTTAACTCTGAGGGCGAGTGTGCAACTCTAACCC	6016
QY	2705	CSTGTCCTTATACCAGCTGAGCTGAGCTGAGCTTATAGCTGAGG	2764	QY	3785	YCNATCTTCAGAGATGAGAGATAACCTGAGATCAACARTAATTCT	3844
Db	4959	TGTTGTTCTCTATAGCCAGCTGAGCTGAGCTTATAGCTGAGG	5018	Db	6017	CCAGATTTCTGAGAGATGAGAGATAACCTGAGATCAACAGTAA	6076
QY	2765	GGAGCAGAGGGTTACSTCTGGAGCTTATAGCTGAGG	2824	QY	3845	AAACCTTGCCACTTGAGGAGCTTGTAGAGAAAGACTGAGCTGAG	3964
Db	5019	GGAGCAGAGGGTTACGAGCTGAGCTGAGCTTATAGCTGAGG	5078	Db	6077	AAACCTTGCCACTTGAGGAGCTTGTAGAGAAAGACTGAGCTGAG	6135
QY	2825	TCCCTGACTCTCAATCTGCTTGTGAGACTTGTGCTTGTAGG	2884	QY	3965	TCCTATGAGAAUTCCCTGTGAGAAAGACTGAGCTGAGCTGAG	3904
Db	5079	TCCCTGACTCTCAATCTGCTTGTGAGACTTGTGCTTGTAGG	5138	Db	6116	TCCTATGAGAAUTCCCTGTGAGAAAGACTGAGCTGAGCTGAG	6195
QY	2885	CTGACTCTTACCCAGGGTCAGGATAGGCCCTCTATGAGG	2944	QY	6077	AAACCTTGCCACTTGAGGAGCTTGTAGAGAAAGACTGAGCTGAG	6255
Db	5119	CTGACTCTTACCCAGGGTCAGGATAGGCCCTCTATGAGG	5198	Db	6136	TCCTATGAGAAUTCCCTGTGAGAAAGACTGAGCTGAGCTGAG	6255
QY	2945	CCAGACTCTGAGCARTYMTCATACTCTGGIACACTGTGCTTGTAG	3004	QY	6256	GCAGACTCTGAGCAGGGAGCTTGTAGGTTGAGSCTTCTGACTGAG	6315
Db	5199	CCAGACTCTGAGCAGGGTCAGGATAGGCCCTCTATGAGG	5256	QY	6256	GCAGACTCTGAGCAGGGAGCTTGTAGGTTGAGSCTTCTGACTGAG	6315
QY	3005	ACTTTTTCGCGYRTTCAGAACCTGTGCCCCATAGCCACCCAGC	3064	Db	6316	TATCACGACTTAACTGAGAACTTGAGGAGCTGAGCTGAGCTGAG	6375
Db	5257	ACTTTTTCGCGCCATTCAAGAACCTGTGCCCCATAGCCACCCAGC	5316	QY	6316	TATCACGACTTAACTGAGAACTTGAGGAGCTGAGCTGAGCTGAG	6375
QY	3065	CCTGCGYACCTGTCAGWGTTCTCAACASBARGCTCARCTCTGAGG	3124	Db	6415	AGGAAGCTCTAACCTCYAGRGAGACCTATCAGCAGGAGCTGAG	4204
Db	5317	CCTGCGTACTGTGCTACTGTGCTACCAACCAAGGCTCAGCAGC	5376	Db	6375	6375	6375
QY	3125	TAATACTCTGGRCTAATATCCAAAGGACCCAGGGCTCAAGGAG	3184	QY	6415	TATCACGACTTAACTGAGAACTTGAGGAGCTGAGCTGAGCTGAG	6415
Db	5377	---TACTTGGCTAAATATCCAAAGGACCCAGGGCTCAAGGAG	5432	Db	6436	TTATGCTGCTACGAGAACCTAAAGGAGCTGAGGAGCTGAGCTGAG	6495
QY	3185	GCCATATCTGCTTATCTCTATCYCRAAACCTAAAGCACTAAGR	3244	QY	6436	TTATGCTGCTACGAGAACCTAAAGGAGCTGAGGAGCTGAGCTGAG	6495
Db	5433	GCCTATACCTGCTTACCTCTTCCAAAGAACCTAAAGCACTAAGG	5492	QY	6496	TTATGCTGCTACGAGAACCTAAAGGAGCTGAGGAGCTGAGCTGAG	6555
QY	3245	AAYGGYTCTGCGCAAATGGATTCCTGGTGTGCRATAAGCCAGG	3304	Db	6496	TTATGCTGCTACGAGAACCTAAAGGAGCTGAGGAGCTGAGCTGAG	6555
Db	5493	AATAGTTCTGCGCAAATGGATTCCTGGTGTGCRATAAGCCAGG	5551	QY	6555	GGAAAGRAAGGAAATASAAGGAAATGCTTAACTTCTTGTAGG	6615
QY	3305	CASTATTAGGAACCTCAGAACCTTACCTTGTAGTGGAGYMC	3364	Db	6555	GGCAGACCTCTCATGAAATGCTTAACTTCTTGTAGG	6615
Db	5552	CACTATTAGGAACCTCAGAACCTTACCTTGTAGTGGAG	5611	QY	6555	GGCAGACCTCTCATGAAATGCTTAACTTCTTGTAGG	6615
QY	3365	AAGTGCTTCCAGGCCCTTAAGAGGGCTTAACCAAGGCC	3424	Db	6616	CGGAAACCAAGCCAGTACTCAGCAGGAGAACAGATGGGAACTC	6675
Db	5612	AAGTGCTTCCAGGCCCTTAAGAGGGCTTAACCAAGGCC	5657	QY	6616	CGGAAACCAAGCCAGTACTCAGCAGGAGAACAGATGGGAACTC	6675
QY	3425	AACGGGCAAGACTTSTTATYRTCAAGAAAAGAGRAVAGCT	3484	Db	6676	TTTCTCCCTCGGAGCTGAGCTGAGGAGAACAGATGGGAACTC	6735
Db	5658	AACGGGCAAGACTTSTTATYRTCAAGAAAAGAGRAVAGCT	5716	QY	6676	TTTCTCCCTCGGAGCTGAGCTGAGGAGAACAGATGGGAACTC	6735
QY	3485	TTACACGRCRAGGAYAGCTGCAACCTGRCYCTGAGTGTAG	3544	Db	6796	TCAGATGGCCAATCTTACTGAGCTGAGGAGAACAGATGGGAA	6855
Db	5717	TTACACGRCRAGGAYAGCTGCAACCTGRCYCTGAGTGTAG	5776	QY	6796	TCAGATGGCCAATCTTACTGAGCTGAGGAGAACAGATGGGAA	6855
QY	3545	TAGTGCAAAAGGGTGTGCTGAGTGTAGTGTAGTGTAGT	3604	QY	6824	CAGGGCTGCTGAGCTGAGGAGAACAGATGGGAACTCCTGCTG	4683
Db	5777	TAGTGCAAAAGGGTGTGCTGAGTGTAGTGTAGTGTAGT	5836	QY	6824	CAGGGCTGCTGAGCTGAGGAGAACAGATGGGAACTCCTGCTG	4683
QY	3605	CTGAGCAGTAATATAGGAGAGATCTACTGAGTGTAGTGTAG	3664	Db	6856	CAGGGCTGCTGAGCTGAGGAGAACAGATGGGAACTCCTGCTG	6915
Db	5837	CTGAGCAGTAATATAGGAGAGATCTACTGAGTGTAGTGTAG	5896	QY	6856	CAGGGCTGCTGAGCTGAGGAGAACAGATGGGAACTCCTGCTG	6915
QY	4744	AAACCTCAGGGATTTCAGTACTAGTCTGGTARACTTCACGG	4803	QY	6884	ARACAAACAGGGCAATACCTGAGAACAGATGGGAACTCCTGCTG	4743
Db				QY	6916	AGACAAAGGGCAATACCTGAGAACAGATGGGAACTCCTGCTG	6975

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="19"
 /isolation_source="PBMC of individual 79"
 /db_xref="Taxon:9506"
 /chromosome="7"
 /map="7q21-q22"
 /sex="male"
 /cell_type="PBMC"
 /note="caucasian
 isolated by PCR; ERVWE1 locus, allele B
 endogenous virus: HERV-W"
 1. .780
 join(248 .811 .2887 .9813)
 /note="putative mRNA transcript 1; gag-pol pseudogene"
 join(248 .811 .7585 .9813)
 /product="envelope glycoprotein"
 /note="putative mRNA transcript 2"
 join(248 .811 .9248 .9813)
 /note="putative mRNA transcript 3; no coding region"
 7812. .9428
 /note="syncytin"
 /codon_start=1
 /product="envelope glycoprotein"
 /protein_id="AAH68164.1"
 /db_xref="GI:37544406"
 /translation="MFLHPLFVLLPSFLTAPPPRCMTSSSPYOEFLWRMOPR
 GMDAHSYRSKGTPTFHMPRNCHSRLTCMHNTHWIGKINPSCPGGLV
 TYCWTFTOTMSDGGOVDRREKVEKPYGKIDSKLKHETLT
 HIRLVLSPFTNTGHEVSQNPNGWICLUNPRPVSIUPEOMNNESTEINTSV
 LVGPMVNLHHTHNTCVCPSNTTNSCIRWPKPVIQVCPGIVFCVGTSV
 YRCNSESSEMFSLSPVPMIYTSDLYSVIKRNPVILPVGIVAGVLGIA
 TEGGTTSTQTYKLSQNDMENVADSVLTDLNSTAVVONRRAIDLTAE
 RGGTCFLGEBCCYYNNQSIVTEKEVKEIRDKIQRRAEERLRTGWLSPWLP
 FGLPLAIIILFGCICLUNPFSSRIBVWKLOPENPKOSKTKYRPLDPRASP
 RSDVNDIKGTPPERBISAQPLLRPNAGGS"
 9487. .10222

ORIGIN
 LTR

Best Local Match 85.0%; Score 6448; DB 9; Length 10222;
 Definition Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVWE1 locus,
 allele B, complete sequence.

ACCESSION AV101585
 VERSION AV101585.1
 KEYWORDS GI:37544405

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens (human)

REFERENCE 1. (bases 1 to 10222)
 AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
 Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
 TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
 in hominoid placental physiology
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 PUBLISHED 14/7/2005
 REFERENCE 2. (bases 1 to 10222)
 AUTHORS Millet,F., Bouton,O. and Oriol,G.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
 CIRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allee
 d'Italie, Lyon 69364 cedex 07, France
 FEATURES source
 1. .10222

6964 CTCCCTTGACCCTGATCTTTRACTCTTGTRACTGTCTTCAGATCGA 7023
 9195 CTCCCTTGACCCTGATCTTACTCTTGTRACTGTCTTCAGATCGA 9254
 QY 7024 GCTGTTAACATACAAATGGGCCAGATCAGTCAGACTAAGATCTACCGAGACCC 7083
 9255 GCTGTTAACATACAAATGGGCCAGATCAGTCAGACTAAGATCTACCGAGACCC 9314
 QY 7084 CTGAGCGGCTGTYAGCCAGCACTGAGCTGAGTTAATGACTCAAGGCAGCCCTCTGAG 7143
 9315 CTGGACCGGCTGCTAGCCAGATCTAGCTGAGCTGAGTTAATGACTCAAGGCAGCCCTCTGAG 9374
 Db 7144 GAAATCTAGCTGACACCTCTACTACAGCCCAATTAGCAGGAGGAGTAGGGGT 7203
 9375 GAATCTAGCTGACACCTCTACTACAGCCCAATTAGCAGGAGGAGTAGGGGT 9434
 QY 7204 SGTCGCCAACTCTCCAAAGCAGCACTGAGCTGAGTTAATGACTCAAGGCAGCCCTCTGAG 7263
 9435 CGTCGCCAACTCTCCAAAGCAGCACTGAGCTGAGTTAATGACTCAAGGCAGCCCTCTGAG 9494
 Db 7264 AGGATAGCTGGATTCTCTAGCTGAGTAACTCYTAGGCTGTTGAGCTGAGTGA 7323
 9495 AGGATAGCTGGATTCTCTAGCTGAGTAACTCYTAGGCTGTTGAGCTGAGTGA 9554
 Db 7324 ACATCCACCTTAACAGGGCTGCACTAGTCACCTGACCATAGAGCT 7383
 9555 ACATCCACCTTAACACGGGCTGCACTAGTCACCTGACCATAGAGCT 9614
 QY 7384 ACTAAATGTTAATGGCAAGACAGAGCAGAGCTAAGAATAGCCATCATTTGCMTC 7443
 9615 ACTAAATGTTAATGGCAAGACAGAGCAGAGCTAAGAATAGCCATCATTTGCMC 9674
 Db 7444 AGAGCACAGGAGGAGACATGATGGGATAATAACCCAGTGTGAGCG 7503
 9675 AGAGCACAGGAGGAGACATGATGGGATAATAACCCAGTGTGAGCG 9734
 Db 7504 CAACCCCTTGGGCCCCCTGTTAGGGAGCTGTTCTGCTGCTACTCT 7563
 9735 CAACCCCTTGGGCCCCCTGTTAGGGAGCTGTTCTGCTGCTACTCT 9794
 QY 7564 ATTAATCTGCARCTGCR 7582
 Db 9795 ATTAATCTGCARCTGCR 9813

RESULT 5

AV101585
 LOCUS AV101585 10222 bp DNA linear PRI 11-FEB-2004
 DEFINITION Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVWE1 locus,
 allele B, complete sequence.

ACCESSION AV101585
 VERSION AV101585.1
 KEYWORDS GI:37544405

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens (human)

REFERENCE 1. (bases 1 to 10222)
 AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
 Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
 TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
 in hominoid placental physiology
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 PUBLISHED 14/7/2005
 REFERENCE 2. (bases 1 to 10222)
 AUTHORS Millet,F., Bouton,O. and Oriol,G.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
 CIRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allee
 d'Italie, Lyon 69364 cedex 07, France
 FEATURES source
 1. .10222

QY 3186 CCAGCATCCGAGCTCTCCCAACTTAAGGACCCCTCAACCAAATGGCTCAA 3245
 Db 3126 AATTAGTAAAGTGTGTTAGCTTCAAGGAGCTTCAGCTACTCTCTATC 3185
 QY 906 CCAGCATCCGAGCTCTCCCAACTTAAGGACCCCTCAACCAAATGGCTCAA 965
 Db 966 AAGGAGTAGACAAAGGGTAACAGTGACCAAGGTCGCAATTCTCCAAATTATGA 1025
 3246 AGGAGTAGACAAAGGGTAACAGTGACCAAGGTCGCAATTCTCCAAATTATGA 3305
 QY 1026 CCCCTCCCAGCACTGGGAGAAGGAATCGGCCAGCAGAGTCATGCTTTCY 1085

Db	3306	CCCTT- CCAAGCAGTGGGAGGAAGAGAATCGGCCAGCCAGTCATGGCTTC 3364	QY	2165	TGTCCCCGACACTGTCCTCAGATCTGTCACTATTCTGAGGGGGTCNTLAGACGGCA 2224
QY	1086	TCTCCCGAGCTTAAGGAAATAAACAGACTTAGGAAATTCAGATAAYCTGTG 1145	Db	4434	TGTCCCCGACACTCTCTCAGATCTGTCACTA TCTGAGGGGGTC -TRAGACGGCA 4491
Db	3365	TCTCCCGAGCTTAAGGAAATAAACAGACTTAGGAAATTCAGATAACCTGAG 3424	QY	2225	GTCACCTAGACTTTTCCAGCCACTAAGTGTAGCACTGGGACCTTATCTTCAC 2284
QY	1146	CTATATTGRTGTTACAGGGTTAGGACCAATTCTTGATCTGACATGGAGATA 1205	Db	4492	GTCACCTAGACT-TCTCCAGCCACTAAGTGTAGCACTGGGACCTTATCTTCAC 4549
Db	3425	CTATATTGRTGTTACAGGGTTAGGACCAATTCTTGATCTGACATGGAGATA 3483	QY	2285	ATGCTTTCTAATTGCTGAAGCCCACACTCTGTAGGGAGACATCTGCGCT 2344
QY	1206	TGTCATGCTGAAATCAGACACTAACCCAAATGAGAGGAGGACCTAATGCGCT 1265	Db	4550	ATGCTTTCTAATTGCTGAAGCCCACACTCTGTAGGGAGACATCTGCGCT 4609
Db	3484	GAGRGTGCGATCTGTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 3543	QY	2345	AAGCAGGGGCATATTACACCTGACACATAGGAGAGAGACACCCGTTGTGT 2404
QY	1266	GAGRGTGCGATCTGTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1325	Db	4610	AAGCAGGGGCATATTACACCTGACACATAGGAGAGAGACACCCGTTGTGT 4668
Db	3544	GAGAGTTGGCATCTGTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 3600	QY	2405	CTTGAGGAGGAAGAATTATCTGTAGTCAGTCTGGGCAACAGGACATAATGAGC 2464
QY	1326	AAAGANATGATTCGCCACAGGCCAGCARGATGTCCTGAGCTASACCTCATGGG 1385	Db	4669	CTTGAGGAGGAAGAATTATCTGTAGTCAGTCTGGGCAACAGGACATAATGAGC 4727
Db	3601	AAAGAGATGATTCGCCACAGGCCAGCAGCAGTCAGTCTGGG-- 3657	QY	2465	GAATGCCCGCTGTCAGTAACTAAGGATTCACITCTCTTCCACAAAGGCA 2524
QY	1386	ACAGAACTGATCTGGGAGATGTCAGTAACTAAGGATTCACITCTCTTCCACAA 1445	Db	4728	GAATGCCCGCTGTCAGTAACTAAGGATTCACITCTCTTCCACAAAGGCA 4787
Db	3658	ACACAGAACTGATGAACT-GGAGATGTTGTCAGTAACTTGTGTGAGA 3716	QY	2525	GTACCCCTCTAGACCCCAAGGCCAACAGGATTCACAAAGATGTGTTAGGACT 2584
QY	1446	GGACTAAAGGAACACTGAGAAATCTGAACTGAACTACTCTAGATGTCACCA 1505	Db	4788	GTACCCCTCTAGACCCCAAGGCCAACAGGACTCAAGATGTTAGGAGCTAAAGC 4847
Db	3717	GGACTAAAGGAACACTGAGAG-AAGCTGATGAACTACTCTAGATGTCACCA 3775	QY	2585	CCAAGGCTTAGTAACACCTGCAACTCTGTGAGTAATCCGTAGTAACTGAGG 2644
QY	1506	GGGGAGGAGGAAGAAATCTGCTGCTGAGCTGAGGAGCATGAGGAGCA 1565	Db	4848	CCAAGGCTTAGTAACACCTGCAACTCTGTGAGTAATCCGTAGTAACTGAGG 4900
Db	3776	-GGGAGGGAGAAATCTCTACTGCTGCTTGTGAGACTGAGCTAGGAGCTGGAGC 3834	QY	2645	CACAGAACCCAGTGGCACGTGGAGGTAGTGTAGGAGCTCAAGGAGCTCAAGA 2704
QY	1566	GTGCTCTCTGTCACCTGATCTCTGAACTGAACTAACTTAAGCTGAGCTTACA 1625	Db	4901	TACAGAACCCACAGACAGTCGGA -GTTAGTGTGAGATCTCAAGTATTCAAT-GAGG 4958
Db	3835	GTGCCCTCTGTCACCTGATCTCTGAACTGAACTAACTTAAGCTGAGCTTACA 3894	QY	2705	CGTGTGCTTTATACCCGCTGACTCTGCGCTTATACTGTGTTGTTCCAAATACAGA 2764
QY	1626	CTCAGTCAGTGCAGACATG -AAAACACTCAAAAGTGCGTAGGCCGGAGCA 1684	Db	4959	TGTGTGTTCTCTATAGCCAGCTGFACTCTGACTGCCCCTTAACTCTGTTCCAAATACAGA 5018
Db	3895	CTCAGTCAGTGCAGACATGAAAGAAAACCTCAAAAGTGCGTAGGCCGGAGCA 3954	QY	2765	GGAGAGAGGTTACASTCTCTGAGCTTMAAGTAGCTGCTGAGTCTGAGTCTG 2824
QY	1685	ACTTGAACACCTATGACTGACTGACTG -AAAACACTCAAAAGTGCGTAGGCC 1744	Db	5019	GGAGAGAGGTTACAGTCCGGACTCTGAGTCTGAGTCTGAGTCTGAGTCTG 5078
Db	3955	ACTTGAACACCTATGACTGACTGACTGACTGTTTATAGAGATGAGCTGGAGC 4014	QY	2825	TCCGTGACTCTCAATCTCTGCTGAGTCTGAGTACTCTAACCCARATCTCACTCAC 2884
QY	1745	AGGGAGAACGGGACAACGGGATAAAAGGACACCCTTGTGCTAATGCGCTAATG 1804	Db	5079	TCTGTGACTCTCAATCTCTGCTGAGTACTCTAACCCACATCTCACTCAC 5138
Db	4015	AGGGAGAACGGGACAACGGGATAAAAGGACACCCTTGTGCTAATGCGCTAATG 4074	QY	2885	CTGGACTCTCAATCTCTGCTGAGTCTGAGTACTCTAACCCACATCTCACTCAC 2944
QY	1805	CAAGTGGACTTGGAGGCTGAGGCTGAGGAAAGCTGGCAATGTAATGCGCTAATG 1864	Db	5139	CTGAGACTTACCTTACCCCAAGGGTCAAGGAGTACTCTAACCCACATCTCACTCAC 5138
Db	4075	CAAGTGGACTTGGAGGCTGAGGCTGAGGAAAGCTGGCAATGTAATGCGCTAATG 4134	QY	2945	CCAGAGCTGAGCARTMCATACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGT 3004
QY	1865	GCTTGCTCTGAGTGGCTCAGAACGACACTTAAAGAACATTGTCAGTAGTAG 1924	Db	5199	CCAGAGCTGAGCCTACCTCTCATCTGAGCTGAGCTGAGCTGAGCTGAGT 5256
Db	4135	GCTTGCTCTGAGTGGCTCAGAACGACACTTAAAGAACATTGTCAGTAGTAG 4194	QY	3005	ACTTTCAGCTGCGCTACAGAACCTCTGCTGAGCTGAGCTGAGCTGAGCTGAGT 3064
QY	1925	CCGCCCTCTGTCATGCCCTTATTCAGGGATCATCTGGAGCCCTGCCCCAG 1984	Db	5257	ACTTTCAGCTGCGCTACAGAACCTCTGCTGAGCTGAGCTGAGCTGAGCTGAGT 5316
Db	4195	CCGCCCTCTGTCATGCCCTTATTCAGGGATCATCTGGAGCCCTGCCCCAG 4254	QY	3065	CCTGGCCTACCTGCGCTACAGAACCTCTGCTGAGCTGAGCTGAGCTGAGT 3124
QY	1985	GGACAAAGGTCTTGTAGTGAAGGCCACTAACCGATGATGTCAGCAGCAGGG 2044	Db	5317	CTCTGCTGACTCTGCGCTACATGTCAGCTGAGCTGAGCTGAGCTGAGCTGAGT 5376
Db	4255	GGACAAAGGTCTTGTAGTGAAGGCCACTAACCGATGATGTCAGCAGGG 4314	QY	3125	TAATCTTGTAGTAACTTACAGGACACGGCTGAGGAGATCTGAGGAGATCTG 3184
QY	2045	TGCCCTGGGAGGCCATCCATACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGT 2104	Db	5377	--TACTTAGGCTTAATTCTCAAGGACAGGAGCTGAGGAGCTGAGCTGAGCTGAGT 5312
Db	4315	TGCCCTGGGAGGCCATCCATACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGT 4374	QY	3185	GCCTATCTGGCTATCTCATCTCAAAACCTTAACGACACTGAGRGRRTCTGTGERT 3244
QY	2105	ATCAGGGCCAGGAGGTCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGT 2164	Db	5433	GCCTATCTGGCTATCTCATCTCAAAACCTTAACGACACTGAGGAGATCTGAGCTGAGT 5494

Qy	3245	AAYAGGYTTGCGGAAWATGGATTCCCGGTWGGCRAATAGCCAGSYCATTAWA 3304	Db	6556	GGCAGGACCTTCATAGAATGCTTATAAACAAACCTAGTATAGGTAATCCCTC 6615
Db	5493	AATAGGTTCTGCGGAATATGGATT-CCCGGTATGGCGAAATAGCCAGGCTTAAATA 5551	Qy	4385	CGGGAAACGAGGCCAGTACTCAGCAGGAGAACAGATGGGAACTCACCGAGG-CAG 4443
Qy	3305	CASTATTAGGAACACTCGAAGCAATTCCATTARRAGTAGTGAYMCTGAYR 3364	Db	6616	CGGGAAACGAGGCCAGTACTCAGCAGGAGAACAGATGGGAACTCACCGAGG-CAG 6675
Db	5552	CACTAATTAGGAACACTCGAAGCAATTCCATTARRAGTAGTGAYMCTGAYR 5611	Qy	4444	TTTCTCCCTCGGAGGTTAGCCACTGAGAAGGAAATACTTTCGCTCAACRAT 4503
Qy	3365	AAGTGCTTTCAGGCCCTTAAGAAGAGGCAATTCCATTARRAGTAGTGAYMCTGAYR 3424	Db	6676	TTTCTCCCTCGGAGGTTAGCCACTGAGAAGGAAATACTTTCGCTCAACRAT 6735
Db	5612	AAGGCGTTCCAGGCC-...-TAACCAAGCCAGTITAGTTGCC 557	Qy	4504	CCATGGAATRACTTAAACCTTCACAACCTTCACTTCACTTCACTTCACTTCACTTCAACCTTCAGGCTTGCAACCCCC 4563
Qy	3425	AACGGGCAAGACTTTSYATYRTCAAGAAAAAGAGRAYAGYCTCTGGAGTC 3484	Db	6735	CCATGGAATACTTAAACCTTCACAACCTTCACTTCACTTCACTTCACTTCAACCCCC 6795
Db	5658	AACAGGCAGAAGCTTTCCTCATATGCTCAG-AAAACAGGAATAGCTCTAGGATC 5716	Qy	4564	TCARATGCCAATCATATTACTGACCAAGGCTTCAAAACTTCAGCARATAK 4623
Qy	3485	TTACAGAGTCCRAGGAGAAGCTGCAACYCTGCRYXICCTGASTAGGAALYTGT 3514	Db	6796	TCAGTGGCAATCTTCAACCTTCAGTGGCTTGAGGCTTGAGGAGAATTCAGGATAGT 6955
Db	5717	TTACAGAGTCCRAGGAGAAGCTGCAACYCTGCRYXICCTGASTAGGAALYTGT 5776	Qy	4624	CAAGGCCTGAACTTAAACCTTCATGAGGAACTTCTGCAAGCTTCTGAGGAC 4683
Qy	3545	TAGTGCCAAAGGGTGRCYCATGTTAVGGTAGTGTGTGCACTTGACTAGGTYKTAGM 3604	Db	6856	CAGGCTCTGAGTGTGAGGAAATATCCCTGCCTTATGCCCAGCTCCTCAGG 6915
Db	5777	TAGTGCCAAAGGGTGRCYCATGTTAVGGTAGTGTGTGCACTTGACTAGGTYKTAGM 5836	Qy	4684	ARACAAARAACAGGCCATTACCTGGRBAAACTGGRACATGTTTCAAACTTCAGGATAGT 6976
Qy	3605	CTGAGAGCACTTAATATCAGSERSRAGACTCTACTGTTGCACTTGCACTTAACTG 3664	Db	6916	AGAACAAAGAACAGGCATTAACCTTCATCAAACTTCTACTAGGCTGATAGACCC 6975
Db	5837	CTGAGAGCACTTAATATCAGGAGAAGCTTACTGTTGCACTTAACTG 5896	Qy	4744	AAACCTCAGGATTCTAGTATCTACTGTTCTGGTAACTTTCAAGGGTTGGCARAG 4803
Qy	3665	AVRGCAACTCTACTGCAAAAGGAGCTTGGCTTGCACTGCACTTAACTG 3724	Db	6976	AAACCTCAGGATTCTAGTATCTACTGTTCTGGTAACTTTCAAGGGTTGGCARAG 7035
Db	5897	ATGGCAACTCTACTGCAAAAGGAGCTTGGCTTGCACTGCACTTAACTG 5956	Qy	4804	GCCTTCCCTGTAGGAGAAAGCCCAAGAGGTTAAAGGACTACTTCATGAAATA 4863
Qy	3725	AGGCTTAACTGCAAAAGGAGCTTGGCTTGCACTGCACTTAACTG 3784	Db	7036	GCCTTCCCTGTAGGAGAAAGGCCAGAGGTTAAAGGACTACTTCATGAAATA 7035
Db	5957	AGGCTTAACTGCAAAAGGAGCTTGGCTTGCACTGCACTTAACTG 6016	Qy	4864	ATTCGAGATTCTGCACTTCCCTGAGGTTACAGAGTAAAGGACTACTTCATGAAATA 4923
Qy	3785	YCNCAATTCTTCAGAACTAAGAGATAAAVATAACTGTCAACARTAAATTCTC 3844	Db	7096	ATTCGAGATTCTGCACTTCCCTGAGGTTACAGAGTAAAGGACTACTTCATGAAATA 7155
Db	6017	CCACATTCTTCAGAACTAAGAGATAAAVATAACTGTCAACAGTAATTCTC 6076	Qy	4924	ACGTAACCCAGGGAGATCCAGGGTTAGGTTACAGTATCTACTGCACTTAACTG 4983
Qy	3845	AAACTATGCACTGGGGACCTTAACTGTTCCYCTGACTGCACTTAACTG 3904	Db	7156	ACATTAACCGAGGATATCCCTAGGGTTAGGTTACAGTATCTACTGCACTTAACTG 7215
Db	6077	AAACCTATGCACTGGGGACCTTAACTGTTCCYCTGACTGCACTTAACTG 6135	Qy	4984	AGGCCACAGCTCTGAGGTTGAGAAATGATGATGAAAYACTCAAGGACATCTAAA 5043
Qy	3905	TTGTAACTGTTGAGTTCTTGTAGAAGAAAGGACTTGGCTGAAAGGTTACAGT 3964	Db	7216	AGGCCACAGCTCTGAGGTTGAGAAATGATGATGAAACACTCAAGGACATCTAAA 7275
Db	6135	TGTGTACTGTTGAGTTCTTGTAGAAGAAAGGACTTGGCTGAAAGGTTACAGT 6195	Qy	5044	AACGAAACCCAGGAAACCACCTCACATGGCTGCTGCTGCTGCTGCTGCTG 5103
Qy	3965	GTCACTGATATGGAAATYTTGAAGTAATCCCTCACTCAGGACTAGTGCTAGTR 4024	Db	7276	AACGAAACCCAGGAAACCACCTCACATGGCTGCTGCTGCTGCTGCTGCTG 7335
Db	6196	GTCACTGATATGGAAATYTTGAAGTAATCCCTCACTCAGGACTAGTGCTAGCTA 6255	Qy	5104	ATCGCAACTTCCCAAAGGAGGACTTAGCCATAGCAAGTGTGTTGGAGGCC 5163
Qy	4025	GCAGAACTAAAGCCCTCACTGAGGAAAGGAAATATA 4084	Db	7336	ATCGCAACTTCCCAAAGGAGGACTTAGCCATAGCAAGTGTGTTGGAGGCC 7395
Db	6256	GCAGAACTAAAGCCCTCACTGAGGAAAGGAAATATA 6315	Qy	5164	TCATAACCATGACCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5223
Qy	4085	TATCGAACTTAACTGAGGAAAGGAAATATA 4144	Db	7396	TCATACCATGACCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7455
Db	6316	TATCGAACTTAACTGAGGAAAGGAAATATA 6375	Qy	5224	CCTTAGCCAAATACAGTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5283
Qy	4145	AGGGAACTCTACTTCYGAAGCACCTTCAGGAAACCTATTGAGRATIA 4204	Db	7456	CCTTAGCCAAATACAGTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7515
Db	6376	AGGGAACTCTACTTCYGAAGCACCTTCAGGAAACCTTCAGGAACTTAAATA 6435	Qy	5284	GAAGAGAACTTCCACCGWWTGCTGACTGCTGACTGCTGCTGCTGCTGCTGCTG 5343
Qy	4205	TTATGGCTGAGAACTCTGAGGAGGTTACACTGCGGGCTCATGAA 4264	Db	7516	GAAGAGAACTTCCACCGWWTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 7575
Db	6436	TTATGGCTGAGAACTCTGAGGAGGTTACACTGCGGGCTCATGAA 6495	Qy	5344	CATCCCTAGTACATCTGGAGGACCTTACCTGCTGCTGCTGCTGCTGCTGCTG 5403
Qy	4265	GGAAAGRAAGGAATASAGRGAAAGCAKATATTGAGCMMAAGAGCTGCA 4324	Db	7576	CATCCCTAGTACATCTGGAGGACCTTACCTGCTGCTGCTGCTGCTGCTGCTG 7635
Db	6496	GGAAAGGAAGGAAATAGGAGAGACTGCGCAAGCAGTATGAGGAAAGGCTGCA 6555	Qy	5404	TTAATGGTGGAGGACTCTGGAGTACATGAGCTGAGTAACTCCCTC 4384

RESULT	7	AX329572	AX329572	Sequence 81 from Patent WO1994629.	56093 bp	DNA	linear	PAT 09-JAN-2002
REFERENCE		LOCUS	ACCESSION	DEFINITION				
AUTHORS	Young, P.B., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.	KEYWORDS	AX329572	VERBOSITY	1	GI:18102550		
TITLE	Cancer gene determination and therapeutic screening using signature gene sets	PATENT	WO 0194629-A	WO1994629.	13-DEC-2001.	JOURNAL	Avalon Pharmaceutical (US)	Location/Qualifiers
FEATURES	1. 56093	SOURCE	/organism="Homo sapiens"	/mol_type="unassigned DNA"	/db_xref="taxon:9606"	ORIGIN		
Query Match Similarity	85.0%	Score	6448	DB	6	Length	56093;	
Matches	6681;	Conservative	95.7%	Pred.	No.	0;		
		Mismatches	195;	Indels	50;	Gaps	27;	
Qy	606	GATGSGAACCTTCCCGGAGACAAAACGCCCTTAGAGCTATCTGAGAATTGGAA	665	Db	30953	GATGGGAAACGTTCCCGGAGACAAAACGCCCTTAGAGCTATCTGAGAATTGGAA	31012	Qy
Qy	666	MCATTGACCTCAGCACTAGAAGAAGACTATTTCTTGACTGCCTCG	725	Db	31013	CCAAATTGACCTCAGCACTAGAAGAAGACTATTTCTTGACTGCCTCG	31072	Qy
Qy	726	GCACCTCTGAGGAAGTATAATTACACATCTACAGCTAGACGTTATTCCTGAGCTTGTGAA	785	Db	31073	GCACCTCTGAGGAAGTATAATTACACATCTACAGCTAGACGTTATTCCTGAGCTTGTGAA	31132	Db
Qy	786	AAGGAAATGGTGGAGTCGCTAACTTCTTCATAGAGACAATC	845	Db	31133	AAGGAAATGGTGGAGTCGCTAACTTCTTCATAGAGACAATC	31192	Qy
Qy	846	ATTTGTAAGTGTATTAGCCCTACCGAAACCTTGAGACTCTACCCATAC	905	Db	31193	ATTTGTAAGTGTATTAGCCCTACCGAAACCTTGAGACTCTACCCATAC	31252	Qy
Qy	906	CCAGCATCCCACCTCTCCCCAMYATAAGGACCCCCCTCAACCAAATGGTC	965	Db	31253	CCAGCATCCCACCTCTCCCCAMYATAAGGACCCCCCTCAACCAAATGGTC	31312	Qy
Db	31253	CCAGCATCCCACCTCTCCCCAMYATAAGGACCCCCCTCAACCAAATGGTC	31312	Qy	966	AAGGAGATAGACAAAGGTAAACAGTACCGAAAGGTGGCAATATCCCA	1025	Db
Db	31313	AAGGAGATAGACAAAGGTAAACAGTACCGAAAGGTGGCAATATCCCA	31372	Qy	1086	TCTCCAGCTAAAGGAACTTCCACCTTCCACCTTCCACCTTCCACCTT	1085	Db
Db	31373	CCCTT-CGAGCTGGAGGAGAGATTCGCCAGCTGCGCTT	31373	Qy	1146	CTATATGTTGTTTACAGGAGCTTCTGAGCTGAGGAGATATA	1145	Db
Db	31432	CTCCCGAGCTAAAGGAACTTCCACCTTCCACCTTCCACCTTCCACCTT	31432	Qy	1206	TGTACTGCTTAATCAGACACTACCCAACTAGAGAGAGCCACCATCTGAGCT	1265	Db
Db	31432	TGTACTGCTTAATCAGACACTACCCAACTAGAGAGAGCCACCATCTGAGCT	31432	Qy	1206	TGTACTGCTTAATCAGACACTACCCAACTAGAGAGAGCCACCATCTGAGCT	1265	Db
Db	31611	GAGAGTTGGCGATCTCTGAGTATGCTGAGCTGAGCTGAGCTGAGCTGAG	31611	Qy	1266	GAGRGTTCGGCATCTCTGAGTATGCTGAGCTGAGCTGAGCTGAG	1325	Db
Db	31668	AAAGAGATGATCCACAGGCCAGCAGTCAGTCCAGTCAGTCAGTCAGTCAG	31668	Qy	1326	AAAGAGATGATCCACAGGCCAGCAGTCAGTCCAGTCAGTCAGTCAG	1385	Db
Db	31784	GGACTTAACTGGAGAATCTGAGAATCTGAGAATCTGAGAATCTGAG	31784	Qy	1386	ACAGAATCAGTAACCTGGAGATGTTGCTGAGCAGCATTTGCTAATG	1445	Db
Db	31725	ACACAGGATCAGACAT-GGAGATGGTGTGAGCACATTCAGTCTGAG	31725	Qy	1446	GGACTTAACTGGAGAATCTGAGAATCTGAGAATCTGAGAATCTGAG	1505	Db
Db	31843	-GGGAGGAGAATCCACTGCTTCTGAGAGACTAAGGGCATGAGAAC	31843	Qy	1506	GGGGAGGGAAATCTGAGAATCTGAGAATCTGAGAATCTGAGAATCTGAG	1565	Db
Db	31902	GTGCCCTCTCTGACTGACTCTGAGCTGAGCCACATTACTTAAAGC	31902	Qy	1566	GTGCCCTCTCTGACTGACTCTGAGCTGAGCCACATTACTTAAAGC	1625	Db
Db	31962	CTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG	31962	Qy	1626	CTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG	1684	Db
Db	32022	ACTTGAACCTTGTAACTTGAACCTTGTAACTTGAACCTTGTAACTT	32022	Qy	1685	ACTTGAACCTTGTAACTTGAACCTTGTAACTTGAACCTTGTAACTT	1744	Db
Db	32081	AGGCAGAACGAGAACGAGGATTAAGAAAGGCCACCCCTTAGCTGAC	32081	Qy	1745	AGGCAGAACGAGAACGAGGATTAAGAAAGGCCACCCCTTAGCTGAC	1804	Db
Db	32082	AGGCAGAACGAGAACGAGGATTAAGAAAGGCCACCCCTTAGCTGAC	32082	Qy	1805	CAACTGACTTGTGAGCTCTGGAAAGTGGCAATTTGAGCTGCTTAA	1864	Db
Db	32202	GCTGTTCCAGTGGCTCACAGGACACTCGGTTTAAATAGAGATCAG	32202	Qy	1865	GCTGTTCCAGTGGCTCACAGGACACTTGAGCTGAGCTGCTTAA	1924	Db
Db	32412	CAAGTGGACTTTGGAGCTCTGGAGCTTGTGAGAAAGGGCAATTTG	32412	Qy	1925	CGGCCCTCTCCATGCCCTTATTTCAAGGAAATCTCGAGGCCACTG	1984	Db
Db	32262	CGGCCCTCTCCATGCCCTTATTTCAAGGAAATCTCGAGGCCACTG	32262	Qy	1985	CGACAAAGGTTTGTGAGCTGAGGCCACTTACCGAGTATGCTGAG	2044	Db
Db	32322	GGCAAAAGGCTCTGAGCTGAGCTGAGCCACTAACAGATGATCCAG	32322	Qy	2045	TGCCTGGCGAGGCCATCCATCACCTTCAAGGCCAGGGCTGGGT	2104	Db
Db	32382	TGCCTGGCGAGGCCATCCATCACCTTCAAGGCCAGGGCTGGGT	32382	Qy	2105	ATGGAGGCCAGGAAGTGTCTCTGGACTCTGGGGCTCTAGCTTAC	2164	Db
Db	32442	ATGGAGGCCAGGAAGTGTCTCTGGACTCTGGGGCTCTAGCTTAC	32442	Qy	2165	TGTCCGGAGACTCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	2224	Db
Db	32501	TGTCCGGAGACTCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	32501	Qy	2225	GTCACTGAGACTCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	2284	Db
Db	32559	GTCACTGAGACTCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	32559	Qy	2285	ATGCTTCTAATGCTGAGAGACATCTGAGAGACATCTGAG	2344	Db
Db	32617	ATGCTTCTAATGCTGAGAGACATCTGAGAGACATCTGAG	32617	Qy	23617	ATGCTTCTAATGCTGAGAGACATCTGAGAGACATCTGAG	32676	Db

QY	2345	AGCGGGGCAATTATCACCTGACCTAGGAGGAGCACCGGTTGTGNNCCCTG	2404	Db
Db	32677	AAGCAGGGCAATTATCACCTGACCTAGGAGGAGCACCGGTTGTGNNCCCTG	32735	Db
QY	2405	CTTGGAGAAGGATTAACTCTGAGTCAGTCGGCAACAGAGGACATATGGAGGCCAA	2464	Db
Db	32795	GAATGCCGCTGTGTCAGTAACTAACAGGATCTCTGAGTCGGCAACAGAGGACATATGGAGGCCAA	32794	QY
QY	2525	GTACCCCTAGACCCAAAGGCCAACAGGATCCAAGATGTTAGGAGCATTTGGAGGAG-CAA	2584	Db
Db	32855	GTACCCCTAGACCCAAAGGCCAACAGGATCCAAGATGTTAGGAGCATTTGGAGGAG-CAA	32914	QY
QY	2585	CCAGGCTTAGAAACATGATACTCCCTGCGAGTAATTCGTTAGGGATTAAAGC	2644	Db
Db	32915	CCAGGCTTAGAAACATGATACTCCCTGCGAGTAATTCGTTAGGGATTAAAGC	32967	QY
QY	2645	CACACAAACCAGTGACAGTGAGGAGTTAGGCAAGATCTCAGGATTATGAGGG	2704	Db
Db	32968	TACAGAAACCAACAGACAGAGCTGGA-GGTTAGTGCAAGATCTCAGGATTATGAT-GAGG	33025	QY
QY	2705	CGTTGCTCTTATACCCAGGTGTTACTAGCCTTATACTGTGMYTCCCAATACCGA	2764	Db
Db	33026	TGTGTTCTCTATAGCCAGGTGTTACTAGCCTTATACTGTGMYTCCCAATACCGA	33085	QY
QY	2765	GGAAACAGAGAGTGTACASPCCT3ACCTTMAGATGCCTCTTGCACTCCGTACA	2824	Db
Db	33086	GGAAACAGAGAGTGTACASPCCT3ACCTTMAGATGCCTCTTGCACTCCGTACA	33145	QY
QY	2825	TCCGACTCTCAATTCTGTTGCCTTGAGATACTCAACCCARCATCTCACTCAC	2884	Db
Db	33146	TCCGACTCTCAATTCTGTTGCCTTGAGATACTCAACCCARCATCTCACTCAC	33205	QY
QY	2885	CTGGACTTTTACCCAAAGGGTTAGGGATAGTYCCCATCTATTGGCCAGGATTAGC	2944	Db
Db	33205	CTGGACTTTTACCCAAAGGGTTAGGGATAGTYCCCATCTATTGGCCAGGATTAGC	33265	QY
QY	2945	CCRAAGACTTGTAGYCARTYMTCACTACTGGACACTCTGTCTCTGTTAKGGAGATT	3004	Db
Db	33266	CCRAAGACTTGTAGYCATTCTCACTCTGGACA-C-CTGTCTCTGTTAKGGAGATT	33323	QY
QY	3005	ACTTTRGGYGCYPTTCAGRAACCTGTGCACTAGGCCACCAAGCRCTTMMATT	3064	Db
Db	33324	ACTTTRGGYGCYCTTCAGRAACCTGTGCACTAGGCCACCAAGCRCTTMMATT	33333	QY
QY	3065	CCTCGCYACCTGTGCTACAGWGGTTCACAGAACASDARGCTACRGTCTGTGACAGCAGGT	3124	Db
Db	33384	CCTCGYACCTGTGCTACAGWGGTTCACAGAACASDARGCTACRGTCTGTGACAGCAGGT	33443	QY
QY	3125	TAACTACTTAGRCRCAARATTCCRAAGGCCACRGCCCTCAAGGAGGAAATTC	3184	Db
Db	33444	---TACTAGGGCTAAATTTCACAGGCCACRGCCCTCAAGGAGGAAATTC	33449	QY
QY	3185	GCCCTTACTCGCTTATCCATCGAACCTAACAGCACTAACGRRRTCTTGCCR	3244	Db
Db	33500	GCCTTACTCGCTTATCCATCGAACCTAACAGCACTAACGRRRTCTTGCCR	33559	QY
QY	3245	AAVAGGYTTCTGCCAAWATGGATTCCCGAGTWTGCGRAATAAGCCAGGYCATAWATA	3304	Db
Db	33560	AAATGGTTCTGCCAAATGGATT-CCAGGTATOGCGAAATAGCCAGGTCTTAA	33618	QY
QY	33619	CACATTAGGAACATCAAGACGCAATCCATTAGGAGTAG	33678	Db
QY	3365	AAGTGGCTTCCAGGCCCTTAAGAAGGCCCTAACCAAGYCCAGTGTAGTGCC	3424	Db
Db	33679	AAGTGGCTTCCAGGCCCTTAAGAAGGCCCTAACCAAGYCCAGTGTAGTGCC	33724	QY
QY	3425	AACRGGGCAAGACTTTSTTYATVTCAGAAAAAACGRAGRAVAGCTTRGGAGTCC	3484	Db
QY	3444	TTTCGCCCCCTGGCAAGGCACTGAGGAGGAAATACCTTGTGCTGCACTAT	4503	Db
Db	34473	TTTCGCCCCCTGGCAAGGCACTGAGGAGGAAATACCTTGTGCTGCACTAT	34802	QY
QY	4504	CAATGGAATTACTTAACCTTCATCAACCTTCACCTAGGATCTGAGTAGACCCA	4563	Db

Db	34803	CCAAATGGAATTACTTAAACCTTCATCAAACTTCATAGGATCGTAGGRCCCA	34852	QY
QY	4564	TCAATGGCCAATTTACTTAACTGACCAGGCTTTCACAACTATCAAGCARATAK	4623	QY
Db	34863	TCAAGTGCCCAATCATTTACTGACCAGGCTTTCACAACTATCAAGCARAGTAGT	34922	Db
Db	4624	CAGGCCGTGAAKGTGCCRARRAATATCCCCTCATGCCAACGCTTCAG	4683	QY
Db	34923	CAGGCCGTGAAKGTGCCRARRAATATCCCCTCATGCCAACGCTTCAG	34982	Db
QY	4684	ARACAAARAACAGGCATTACCCCTTARAARACTGCAACTGATTACCCCTGAGGCC	4743	Db
QY	34983	AGAAARAAAGACAGGCCATTACCCCTGAGAGAACTGCAACTGATTACCCACAGCC	35042	QY
QY	4744	AAACTCAGGATTTAGTACTCATAGTCAGTCTGGGTTGGSCARAG	4803	Db
Db	35043	AAACCTCAGGATTCTAGTACTCATAGTCAGTCTGGGTTGGSCAGAG	35102	QY
QY	4804	GCTTCCCTGTAGGACAGAAGGCCAAGGGATAAAGGCACTGACTGATGAAATA	4863	Db
Db	35103	GCCTTCCCTGTAGGACAGAAGGCCAAGGGATAAAGGCACTGACTGATGAAATA	35152	QY
QY	4864	ATTCAGATTCGGACTTCCCAGGCTTACAGTGACAATAGCCCTGCTTCAGGC	4923	Db
Db	35153	ATTCAGATTCGGACTTCCCAGGCTTACAGTGACAATAGCCCTGCTTCAGGC	35222	QY
QY	4924	ACAGTAACCCAGGGATTCAGGCTTAGTACGATACACTGCGCTGA	4983	Db
Db	35223	ACAGTAACCCAGGGATTCAGGCTTAGTACGATACACTGCGCTGA	35282	QY
QY	4984	AGGCCACAGTCTCAGGGAGGTGCGAGAAATGAGAAACTCHAGGACATCTAA	5043	Db
Db	35283	AGGCACAGTCTCAGGGAGGTGCGAGAAATGAGAAACTCHAGGACATCTAA	35342	QY
QY	5044	AAGCAACCCGGAACCCACTCAGTGAGGAACTGAGAAATGAGAAACTCHAGGACATCTAA	5103	Db
Db	35343	AAGCAACCCGGAACCCACTCAGTGAGGAACTGAGAAATGAGAAACTCHAGGACATCTAA	35402	QY
QY	5104	ATCTGCACTTCCCAAAGCAGGACTAGCCATAGAATGGCTATGGAAAGCC	5163	Db
Db	35403	ATCTGCACTTCCCAAAGCAGGACTAGCCATAGAATGGCTATGGAAAGCC	35462	QY
QY	5164	TTCATACCAATGACCTGTCGTTGACCAAGAGCCAACTGAGCAACTTAGTGAGCATACCT	5223	Db
Db	35463	TTCATACCAATGACCTGTCGTTGACCAAGAGCCAACTGAGCAACTTAGTGAGCATACCT	35522	QY
QY	5224	CCTTACCAATATCAACAGTTAACACATTACAGGAACCTATCCCTGAGAGGG	5283	Db
Db	35523	CCTTACCAATATCAACAGTTAACACATTACAGGAACCTATCCCTGAGAGGG	35582	QY
QY	5284	GAAGAGACTATCACCCWGTGACATGGTATAGTCAGTCCTCYCTTAATCC	5343	Db
Db	35583	GAAGAGACTATCACCCWGTGACATGGTATAGTCAGTCCTCYCTTAATCC	35642	QY
QY	5344	CATCCCTAGATACATCCCTGGAAAGACCTACCCAGTCATTTTAACTCCCAACTGGG	5403	Db
Db	35643	CATCCCTAGATACATCCCTGGAAAGACCTACCCAGTCATTTTAACTCCCAACTGGG	35702	QY
QY	5404	TAACTGGCTGAGTGGACTGAGCTTCAGGAGGAAGTAACATAATCTAATCCCCATG	5463	Db
Db	35703	TAACTGGCTGAGTGGACTGAGCTTCAGGAGGAAGTAACATAATCTAATCCCCATG	35762	QY
QY	5464	CAAGGAACCTGAAATCCAGGAGCAACGCTGATTCCTGAACTCTAGGATT	5523	Db
Db	35763	CAAGGAACCTGAAATCCAGGAGCAACGCTGATTCCTGAACTCTAGGATT	35822	QY
QY	5524	TGGCCCTGCTCTAACACACACAGGAGGAAGTAACATAATCTAATCCCCATG	5583	Db
Db	35823	TGGCCCTGCTCTAACACACACAGGAGGAAGTAACATAATCTAATCCCCATG	35881	QY
QY	5584	GCCTCCCTATCATATTTCCTKTSTGTTTACCCSTTCACTCTACTGCA	5643	Db
Db	35882	GCCTCCCTATCATATTTCCTKTSTGTTTACCCSTTCACTCTACTGCA	35941	QY
QY	5644	CCCCCTCATCCGCTTATCCAGTAGGTCCTTCACAGAGTCATGGAGAT	5703	Db
Db	35942	CCCCCTCATCCGCTTATCCAGTAGGTCCTTCACAGAGTCATGGAGAT	36001	QY
QY	5704	CAGCGTCCGCGAATATGATGCCCATCGTATAGGAGCTTISTAGGAAACCCACC	5763	Db
Db	36002	CAGCTTCCGCGAATATGATGCCCATCGTATAGGAGCTTISTAGGAAACCCACC	36061	QY
QY	5764	TTCAGTGCCGACCCATGCCCGCAACTGCTTACCTGCTACTGCTGATGCT	5823	Db
Db	36062	TTCAGTGCCGACCCATGCCCGCAACTGCTTACCTGCTACTGCTGATGCT	36121	QY
QY	5824	GCATRACTCTTATGGAGGAAATGATTTACCTCTAGTGTGCTGGGGAGCTGAA	5883	Db
Db	36122	GCATRACTCTTATGGAGGAAATGATTTACCTCTAGTGTGCTGGGGAGCTGAA	36181	QY
QY	5884	GTCACTGTGTTGGACTTACTTCACCCAACTGGTGTGATGGGGTGGAGTC	5943	Db
Db	36242	GTCAGCAGAGAAACATGTAAGAAGCTGAGGAACTGCTTACCTGCTGAGTC	36301	QY
QY	6004	ACCTCTAGCCCTACAGAGCTAGATCTCAAACACTGGTAAACCTCGTGTACCCAT	6063	Db
Db	36302	ACCTCTAGCCCTACAGAGCTAGATCTCAAACACTGGTAAACCTCGTGTACCCAT	36361	QY
QY	6064	ACTCGCTGGAGGCTTATTATACACCCTCACTGGCTCATGGGTTGGCCCA	6123	Db
Db	36362	ACTCGCTGGAGGCTTATTATACACCCTCACTGGCTCATGGGTTGGCCCA	36421	QY
QY	6124	AACTCTACTAATGTTGAGCTTATTAACCTGCTTACCTGCTCCTCAGCTTCCAACTCAGGCTCATGGTTOATCCT	6183	Db
Db	36422	AACTCTACTAATGTTGAGCTTATTAACCTGCTCCTCAGCTTCCAACTCAGGCTCATGGTTOATCCT	36481	QY
QY	6184	GTACTGAGCATGGAGCAATTGAGGAACTGAGAAATAACCAACTCGTTAGAGA	6243	Db
Db	36482	GTACTGAGCATGGAGCAATTGAGGAACTGAGAAATAACCAACTCGTTAGAGA	36541	QY
QY	6244	CTCTCTTTCCATSTGGAATAACCATCTCCATCTCAAACTCACTGTTAGAG	6303	Db
Db	36542	CTCTCTTTCCATSTGGAATAACCATCTCCATCTCAAACTCACTGTTAGAG	36601	QY
QY	6304	AATACTACATCACACCAACTCCATGAGGAACTTACCTGAGGAACTTACCTGTTAGAG	6363	Db
Db	36602	AATACTACATCACACCAACTCCATGAGGAACTTACCTGTTAGAG	36661	QY
QY	6364	GTCTGCCTACCCCTAGGAAATTGTCGTTGGACTCCTACCTGATGTTGAT	6423	Db
Db	36662	GTCTGCCTACCCCTAGGAAATTGTCGTTGGACTCCTACCTGATGTTGAT	36721	QY
QY	6424	GTCTGCCTACCCCTAGGAAATTGTCGTTGGACTCCTACCTGATGTTGAT	6483	Db
Db	36722	GTCTGCCTACCCCTAGGAAATTGTCGTTGGACTCCTACCTGATGTTGAT	36781	QY
QY	6484	GAACAGTTTACAGTTGAGCTATCTAAGGCCCGCACAAAGGATCCATTCT	6543	Db
Db	36782	GAACAGTTTACAGTTGAGCTATCTAAGGCCCGCACAAAGGATCCATTCT	36841	QY
QY	6544	GTCTGCCTACCCCTAGGAAATTGTCGTTGGACTCCTACCTGATGTTGAT	6603	Db
Db	36842	GTCTGCCTACCCCTAGGAAATTGTCGTTGGACTCCTACCTGATGTTGAT	36901	QY
QY	6604	ACCTCTACTGCTACTCAAACTACTCTAAGGAGCTTACCTGCTACTGCT	6663	Db
Db	36902	ACCTCTACTGCTACTCAAACTACTCTAAGGAGCTTACCTGCTACTGCT	36961	QY
QY	6664	GGCAGCTCCCTGTCACCTGAGATACTTAACCTCCCTASCAGCAGTAGCTCRA	6723	Db
Db	36962	GGCAGCTCCCTGTCACCTGAGATACTTAACCTCCCTASCAGCAGTAGCTCRA	37021	QY

QY 6724 ATTCGAGAAGCTTGTACTGCTAACCGCAGAAGGGGGACCTGTTATTAGG 6783
 |||||:|||||:|||||:|||||:|||||:
 Db 37022 AATCGAGAAGCTTGTACTGCTAACCGCAGAAGGGGGACCTGTTATTAGG 37081
 QY 6784 GAAGATGCTTATATGTTATCATCCGAATGGTCACTGAGAAAGTTRAAGAATT 6843
 |||||:
 Db 37082 GAAGATGCTTATATGTTATCATCCGAATGGTCACTGAGAAAGTTRAAGAATT 37141
 QY 6844 CSAGATCGATAACAGTAGCAGCAGACTTGAAACCTGGACCTGGCTTCCTC 6903
 |||||:
 Db 37142 CGAGTCGAAACAAAGTAGCAGCAGGAGGCTTCAG 37201
 QY 6904 AGCCRATGGATGCCCTGGATCTCCCTTTAGACCTCTAGCAGTATATGGTA 6963
 |||||:
 Db 37202 AGCCATATGGTGCCCTGGATCTCCCTTTAGACCTCTAGCAGTATATGGTA 37261
 QY 6964 CTCTCTTGTGACCTGTATTTRACTCTTGTTAATTGTCCTTCAGATCGA 7023
 |||||:
 Db 37262 CTCTCTTGTGACCTGTATTTRACTCTTGTTAATTGTCCTTCAGATCGA 37321
 QY 7024 GCTGTRAACTACAAATGGGCCAAGATCAGTCAGTCAGAAGTCAACCGAGCC 7083
 |||||:
 Db 37322 GCTGTAACATACAAATGGGCCAAGATCAGTCAGTCAGAAGTCAACCGAGCC 37381
 QY 7084 CTGACCCGCTGTTAGCCGATCTGATGTTATGACRTCAAGGCACCCCTCAG 7143
 |||||:
 Db 37382 CTGGACCGGCTGCTGAGCCACGATGTTATGACRTCAAGGCACCCCTCAG 37441
 QY 7144 GAATCTCGCTGACAACCTCTACTACCCCCAATTCGAGGAGCAGTAGGGT 7203
 |||||:
 Db 37442 GAATCTCGCTGACAACCTCTACTACCCCCAATTCGAGGAGCAGTAGGGT 37501
 QY 7204 SGTCGGCCACCTCCCACAGACACTTAGGTTCTGAGATGGGGACTGAGAC 7263
 |||||:
 Db 37502 CGTGGCCACCTCCCACAGACACTTAGGTTCTGAGATGGGGACTGAGAC 37561
 QY 7264 AGRACTAGCTGGATTCTCTAGGCTGATGTTAGAATCCCTAGCTAGCTGGAGGTGACC 7323
 |||||:
 Db 37562 AGRACTAGCTGGATTCTCTAGGCTGATGTTAGAATCCCTAGCTAGCTGGAGGTGACC 37621
 QY 7324 ACCTCACCTTAAACACGGGCTGCAACTTGTACCTTACCAATCAGAGCTC 7383
 |||||:
 Db 37622 ACATCCACCTTAAACACGGGCTGCAACTTGTACCTTACCAATCAGAGCTC 37681
 QY 7384 ACTAAATGCTTAATGGAAAGCAGGGGTAAGAAATAGCAATCATYTATGCTG 7443
 |||||:
 Db 37682 ACTAAATGCTTAATGGAAAGCAGGGGTAAGAAATAGCAATCATYTATGCTG 37741
 QY 7444 AGAGCACAGGAGGACATGATCGGATAAACCAAGTTCCAGGGCAACGG 7503
 |||||:
 Db 37742 AGAGCACAGGAGGACATGATCGGATAAACCAAGTTCCAGGGCAACGG 37801
 QY 7504 CAAACCCCTTGTGCTCTCTTGTGAGGCTGCTGTTCACTTATTCATCT 7563
 |||||:
 Db 37802 CAAACCCCTTGTGCTCTCTTGTGAGGCTGCTGTTCACTTATTCATCT 37861
 QY 7564 ATTAATCTGCACRGCR 7582
 |||||:
 Db 37862 ATTAATCTGCACRGCR 37880

RESULT 8
 HSAC000064 HSAC000064 56093 bp DNA linear PRI 11-NOV-1996
 LOCUS Human BAC clone RG083M05 from 7q21-7q22, complete sequence.
 DEFINITION Human BAC clone RG083M05 from 7q21-7q22, complete sequence.
 VERSION 1
 ACCESSION HSAC000064.1 GI:1669369
 KEYWORDS HTG,
 Homo sapiens (human)
 SOURCE Organism
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 56093)

AUTHORS Pauley,A.
 TITLE The sequence of H. sapiens BAC clone RG083M05
 JOURNAL Unpublished (1996)
 2 (bases 1 to 56093)
 REFERENCE Waterton,R.
 AUTHORS Direct Submission
 TITLE Submitted (23-NOV-1996)
 JOURNAL Genome Sequencing Center
 COMMENT Department of Genetics Washington University
 St. Louis, MO 63108, USA
 e-mail: sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping between sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:
 This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 97SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8194-8197 (1992); Kim et al., Genomics 34:213-218 (1996).
 VECTOR: pBelo
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
 The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H_RG083M05; actual end is at 56093 of H_RG083M05

This clone contains STS SWSS1725.

FEATURES source
 location/qualifiers
 1. -56093
 /organism="Homo Sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:5606"
 /chromosome="7"
 /map="7q21-7q22"
 /clone="H_RG083M05"
 /clone_id="CTB-978SK-B"
 /clone_lkb="CTB-978SK-B"
 /complement="CTB-978SK-B"
 /repeat_region complement(838...1131)
 /rpt_family="ALU"
 <1360...16971
 /gene="WIGSC:H_RG083M05.1"
 /join(<1360...1503,4181...4370,4587...4774,6422...6556,
 9483...9547,11631...11773,11864...12021,13131...13296,
 14885...14988,16349...16546,16837...16971)
 /gene="WIGSC:H_RG083M05.1"
 /note="Alphase: strong similarity to peroxisome
 biosynthesis protein PAST (PID:gi172019); coded for by
 human cDNA C04279 (NID:9146750)"
 /codon_start=1
 /protein_id="PAB46346.1"
 /db_xref="GI:1669371"
 /translation="KRLENIQTKEVAFSEAVWMQPSVVLDDDDLAGIAPVPBEH
 SPDABORCLIVLNUVTKNKKDINPKDPTDQLRHTKETSGFVARDTSHRL
 TTIQPAVKLKKRKPFLPANLPQRQGILICPPGPIKLLAGVABERMPFV
 KGPENLSKYGASQAVDIFRQAQKTCIFDFDEEKAIPRQHDNGVTDVNO
 LITOLDVGEGGLYVPLRSPDPLRSLRSPVPLRSLRSPVPLRSLRSPVPLRSLR
 OMHSLSVSEIILVLSDFLADDVLOHVASYTSSTGADJAKLYNAQLEAH
 MUSKMSLIPDSFKNNYRLYCGSYSEBGNOTISDSISOCISAPSMTDQPGV
 KKDOLFSOPVPLUTASOGCQELTOEDQUDLADISIKGRYNSQSGBEDSMQPGV
 KTKLIGSOSHMLNQHTRPSSIDEWDKNAEL"

repeat_region
 /rpt_family="ALU"
 complement(448...5130)

Qy	966	AGGAGATAGACAAAGGTAAACAGTGAACCAAGACTGCCATATTCCCATTATGA	1025	Db	32382		TGCCTGGCAGGCCCATCCATCACCTCAAGGCCCTGGTAGTCACC	32441			
Db	31313	AAGGAGATAGACAAAGGTAAACAGTGAACCAAGACTGCCATATTCCCATTATGA	31372	Qy	2105	ATGAGGCCAGGAGGTGCTCTCGACACTGGCGGCTTCCTAGTCCTACTTC	2164	Db	32442	ATGGAGGCCAGG-AAGTTGTCCTCGACACTGGCGGCTTCCTAGTCCTTC	32500
Qy	1026	CCCTCCGAAGCAGTGGAGGAGAGAATTGCCAACAGGAGTCGCTGCGCTTYY	1085	Db	31373	CCCTCCGAAGCAGTGGAGGAGAGAATTGCCAACAGGAGTCGCTGCGCTTYY	31431	Qy	2165	TGTCGGACACTGTCCTCAGATCTGTCACTATCTGAGGGGCTNTAGACGGCA	2224
Qy	1086	TOTCCAGACTTAAGCAATAAAAAGATCTAGGTTAATCTGAGAATCTGAGTGG	1145	Db	31432	TOTCCAGACTTAAGCAATAAAAAGATCTAGGTTAATCTGAGAATCTGAGTGG	31491	Db	32501	TGTCGGACACTGTCCTCAGATCTGTCACTA-TCTGAGGGGCTC-TANGACGGCA	32558
Qy	1146	CTATATGRTGTTTACAGGTTAGCACTCTTGATCTGACATGGAGATATA	1205	Qy	2225	GTCACTAGATCTTTCERGCCACTAAGTGTAGACTGGGACTTTCTTCAC	2284	Db	32559	GTCACTAGATAC-TCTCCAGCACTAAGTGTAG-ATCTGGGACTTTCTTCAC	32616
Db	31492	CTATATGRTGTTTACAGGTTAGCACTCTTGATCTGACATGGAGATATA-A	31550	Qy	2285	ATGCTTCTTAAATTGTTGAACCCACTACCTGTCAGTGTCACTATCTGAGGGGCTNTAGACGGCA	2344	Db	32617	ATGCTTCTTAAATTGTTGAACCCACTACCTGTCAGTGTCACTATCTGAGGGGCTNTAGACGGCA	32676
Qy	1206	TGTCTATGCTAAATCAGACACTAACCCAAATGAGAAGGTCGACCAATACTGAGCT	1265	Qy	2345	AAGCAGGGCCTTAACTGAGGAGACATCTGTCAGTGTCACTACCTGTCAGTGTCACTA	2404	Db	32677	AAGCAGGGCCTTAACTGAGGAGACATCTGTCAGTGTCACTACCTGTCAGTGTCACTA	32735
Db	31551	TGTCTATGCTAAATCAGACACTAACCCAAATGAGAAGGTCGACCAATACTGAGCT	31610	Qy	2405	CTTGAGGAGGAACTTACCTGAGGAGACATCTGTCAGTGTCACTATGGGAC	2464	Qy	2465	GAATGCCGCTCTGTCAGTAACTAAGGATTCCACTCTTCCCTACAAAGGCA	2524
Db	1266	GAGRGTGTCGATCTGTCAGTGTCACTACCTGTCAGTGTCACTAAGCAGAGG	1325	Db	32736	CTTGAGGAGGAACTTACCTGAGGAGACATCTGTCAGTGTCACTATGGGAC	31724	Db	32795	GAATGCCGCTCTGTCAGTAACTAAGGATTCCACTCTTCCCTACAAAGGCA	32854
Db	31611	GAGRGTGTCGATCTGTCAGTGTCACTACCTGTCAGTGTCACTAAGCAGAGG	31667	Qy	2525	GTACCCCTCAGACCAAGGCCAACAGGATCTGTCAGTGTCACTAAGCAGAGG	2584	Db	32855	GTACCCCTCAGACCAAGGCCAACAGGATCTGTCAGTGTCACTAAGCAGAGG	32914
Qy	1326	AAGANAVATGTTCCACAGCACCAGCAGTTCAGCTASACCTATTGGGAC	1385	Db	31725	ACAGACGATCATGACAT-GGAGATGTCGTCAGTGTCACTATGGTGTCAA	31783	Qy	2585	CCAACTGTTGTAACCATACCTGAGGAGCATCTGTCAGTGTCACTATGGGAG	2644
Db	31668	AAGAGAGATGTTCCACAGCACCAGCAGTTCAGCTACCTGTCAGTGTCACTATGGG	31724	Qy	1446	GACTAAGAAACTASGAGAAARTTAYGATTACTCATGATGTCACCATACAA	1445	Db	32915	CCAACTGTTGTAACCATACCTGAGGAGCATCTGTCAGTGTCACTATGGGAG	32967
Qy	1386	ACAGAAATCAGTAACTGGGAGATGTCAGTGTCACTACCTGTCAGTGTCACTA	31783	Db	31843	GGGAGAGGAAAGAAATCTACTGCGCTCTGGAGACTAGGAGGCA	31901	Qy	2645	CACGACGACCTGAGTGAACGAGTGGGTTAGTGCAGTCAGGATTAATGAGG	2704
Db	1446	GACTAAGAAACTASGAGAAARTTAYGATTACTCATGATGTCACCATACAA	1505	Db	31902	GTGCGCTCTCTCACCTGACTCTCTGAGGSCAACTAATTAAAGGTTAATCA	31961	Db	32968	TACGAGAACCAACAGA-CAGTGA-GTTAGTGCAGATCTCAGGATTAT-GAGG	33021
Db	31784	GGGAGAGGAAAGAAACTCTGAGGAGCATGAGGAAAGGAAACGCTTAAAGGTTAATCA	31842	Qy	1506	GGGAGAGGAAAGAAACTCTGAGGAGCATGAGGAGCATGAGGAG	1565	Db	32915	CCAACTGTTGTAACCATACCTGAGGAGCATCTGTCAGTGTCACTATGGGAG	32967
Qy	1506	GGGAGAGGAAAGAAACTCTGAGGAGCATGAGGAGCATGAGGAG	1565	Db	31843	-GGGAGAGGAAAGAAATCTACTGCGCTCTGGAGACTAGGAGGCA	31901	Qy	1566	GTGCGCTCTCTCACCTGACTCTCTGAGGSCAACTAATTAAAGGTTAATCA	1625
Db	1566	GTGCGCTCTCTCACCTGACTCTCTGAGGSCAACTAATTAAAGGTTAATCA	1625	Db	31902	GTGCGCTCTCTCACCTGACTCTCTGAGGSCAACTAATTAAAGGTTAATCA	31961	Qy	1626	CTCAGTCAGTCAGCAGACATGAGGAGCATGAGGAGCATGAGGAG	1684
Qy	1626	CTCAGTCAGTCAGCAGACATGAGGAGCATGAGGAGCATGAGGAG	1684	Db	31962	CTCAGTCAGTCAGCAGACATGAGGAGCATGAGGAGCATGAGGAG	32021	Db	32082	AGGCGGAACAGAACAAAGGCCAACGGTTAGTCAGTACCTCAGTACCTCAGG	32081
Db	1685	ACTTGAACACCTATTGAACTTGCACCTGAACTTAAAGATGTCACAGGAGG	1744	Qy	1744	AGGCGGAACAGAACAAAGGCCAACGGTTAGTCAGTACCTCAGTACCTCAGG	1804	Qy	2765	GGAAAGGAGGAGGTTAACSTCCTGAGGACTCTGGCTCTGCACTCCCTGACA	2824
Db	32022	ACTTGAACACCTATTGAACTTGCACCTGAACTTAAAGATGTCACAGGAGG	32081	Db	32082	AGGCGGAACAGAACAAAGGCCAACGGTTAGTCAGTACCTCAGTACCTCAGG	32081	Qy	2805	GGAAAGGAGGAGGTTAACSTCCTGAGGACTCTGGCTCTGCACTCCCTGACA	33145
Qy	1745	AGGCGGAACAGAACAAAGGCCAACGGTTAGTCAGTACCTCAGTACCTCAGG	1804	Db	32082	AGGCGGAACAGAACAAAGGCCAACGGTTAGTCAGTACCTCAGTACCTCAGG	32081	Qy	2825	TCCTCACTCTCAATTCTGTTGCTCTGAGATCTCAACCCARCATCTCACTAC	2884
Db	32142	CTGAGGACTCTTGGGACCTCTCGAGAACGGAAAGGAAAGCTGCGCAATGTCAGTAA	32201	Db	32082	AGGCGGAACAGAACAAAGGCCAACGGTTAGTCAGTACCTCAGTACCTCAGG	32081	Db	33146	TCCTCACTCTCAATTCTGTTGCTCTGAGATCTCAACCCACACATCTCACTAC	33205
Qy	1865	CCTTGCCTCTCATGCCATTTCAGGAAACTTAAGATGTCACAGGAGTAG	1924	Qy	2885	CTGGACTTTTACCCAAAGGTTCAGGATAGYCCCATCTATTGCGGGCATTC	2944	Db	32202	CTTGTACCTCTGAGCACTCTGAGGAGCATGAGGAGCATGAGGAG	32261
Qy	1805	CTAGTGGACTCTTGGGACCTCTCGAGAACGGAAAGGAAAGCTGCGCAATGTCAGTAA	1864	Db	32202	CTTGTACCTCTGAGCACTCTGAGGAGCATGAGGAGCATGAGGAG	32261	Qy	3005	ACTTURGCGCYRTTCAAGAACCTTGTCCATCAAGCCACCAAGRCCTTMAATT	3064
Db	32142	CTGAGGACTCTTGGGACCTCTCGAGAACGGAAAGGAAAGCTGCGCAATGTCAGTAA	32201	Db	32202	CTTGTACCTCTGAGCACTCTGAGGAGCATGAGGAGCATGAGGAG	32261	Qy	2945	CCAAGACTGAGYCATYMCATACCTGGACACTCTGTCCTCTCTGAGGATCT	3004
Qy	1925	CGGCCCTCTCATGCCATTTCAGGAAACTTAAGATGTCACAGGAGTAG	1984	Db	32202	CTTGTACCTCTGAGCACTCTGAGGAGCATGAGGAGCATGAGGAG	32261	Db	32266	CTTGTACCTCTGAGCACTCTGAGGAGCATGAGGAGCATGAGGAG	33323
Db	32262	CTGAGGACTCTTGGGACCTCTCGAGAACGGAAAGGAAAGCTGCGCAATGTCAGTAA	32321	Qy	3065	CCTCGCYACCTGTCGCTCAAGGGTTCMAASRARGTCARCTGTCACAGG	3124	Db	33384	CCTCGCYACCTGTCGCTCAAGGGTTCMAASRARGTCARCTGTCACAGG	33443
Qy	1985	GAGCAAAAGGTGTTGGAGCTAACAGGAGCATGAGGAGCATGAGGAG	2044	Db	32322	GAGCAAAAGGTGTTGGAGCTAACAGGAGCATGAGGAGCATGAGGAG	32381	Qy	3125	TAATFACTTAAGCTAAATTCAAGAACGACGAGGCCCTGGGTATGCTGACC	2104
Qy	2045	TGCTCTGGGCAAGGCCATCCCATGCCATCACCTCACAGAGGCCCTGGGTATGCTGACC	2104	Qy	3184	TAATFACTTAAGCTAAATTCAAGAACGACGAGGCCCTGGGTATGCTGACC	3184				

QY	5344	CATCCCTAGTACATCCTGSSAAGGACCTACCCAGTCATTATTAACCCAACTGCGG	5403	Db	36722	TCAGATCTAGGCTTCTCTCATCTTAGGCCCTATGACCCTCACRT	36781
Db	35643	CATCCCTAGTACATCCTGGGAGGACCTACCCAGTCATTATTAACCCAACTGCGG	35702	QY	6484	GAACAGAATTATCAGTTATGTCATCTAAGGCCCGAACAAAGACTAACATCCT	6543
QY	5404	TAAAGTGGGGAGTGGGAGCTTGATCTACATCTCACITCTAGCAATCTGGATACTGC	5463	Db	36782	GAACAGAATTATCAGTTATGTCATCTAAGGCCCGAACAAAGACTAACATCCT	36841
Db	35703	TAAAGTGGGGAGTGGGAGCTTGATCTACATCTCACITCTAGCAATCTGGATACTGC	35762	QY	6544	CCTTGTATAGGAGGAGGAGCTGGACTAGTGTCAATCTGGATACTGC	6603
QY	5464	CAAGGAACCTGAATACTCCAGGAGAACGCTAGCTACATCACACTTGAGTCATA	5523	Db	36842	CCTTGTATAGGAGGAGGAGCTGGACTAGTGTCAATCTGGATACTGC	36901
Db	35763	CAAGGAACCTGAATACTCCAGGAGAACGCTAGCTACATCACACTTGAGTCATA	3522	QY	6604	ACCTCTACTCAGTCATCACAACTATCTCAGAACAATAGGGACATGAGACGGTC	6663
QY	5524	TGCGCTGCTTCAAC	5583	Db	36902	ACCTCTACTCAGTCATCACAACTATCTCAGAACAATAGGGACATGAGACGGTC	36961
Db	35823	TGCGCTGCTTCAAC	35881	QY	6654	CCTTGTATAGGAGGAGGAGCTGGACTAGTGTCAATCTGGATACTGC	6723
QY	5584	GSCCTCCCTATCATATTCTCTKASTGSTSTTYCCTSTTCACTCTGCA	5643	Db	36962	GGCCACTCTCATTCATCTGAGCTTACCTCTTACCTCTTACCTCTGCA	37021
Db	35882	GGCCACTCTCATTCATCTGAGCTTACCTCTTACCTCTTACCTCTGCA	35941	QY	6724	AATCGAGAGCTTAGACTYCTACCTGCTTACCTCTTACCTCTGCA	6783
QY	5644	CCCCCTCATGCCGTGTATGACCGTAGTGTCCCTYACOMAGAGTTCTATGGAGA	5703	Db	37022	AATCGAGAGCTTAGACTYCTACCTGCTTACCTCTGCA	37081
Db	35942	CCCCCTCATGCCGTGTATGACCGTAGTGTCCCTTACCTAGGACTTAGGAGA	36001	QY	6784	GACGATCTGTTATATCACTGGGATGACTGACTGACTGAGAAGTRAGAATT	6843
QY	5704	CAGCTCCCGAANTATTGTGCCCCATCTCTAGGAGCTTSTAAGGACCCACC	5763	Db	37082	GAAGATCTGTTATATCACTGGGAGTAACTGAGTAACTGAGTAACTGAGTAA	37141
Db	36002	CAGGTCCCAGAATTATGTGCCCCATCTCTAGGAGCTTSTAAGGACCCACC	36061	QY	6844	CSAGATGAGATACAGTAKAGGAGRAAGCTGTTGAAACACTGAGCTTGTCA	6903
QY	5764	TCTCTGCCACACCATATGCCCGCAACTGCACTACTCTGACTTGTCA	5823	Db	37142	CGAGATCGATACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC	37201
Db	36062	TCTACTGCCACACCATATGCCCGCAACTGCACTACTCTGACTTGTCA	36121	QY	6904	ACCCRATGGCCTGGATTCTCCCTCTTAGGACTCTAGCAGCTTAATTTGCA	6963
QY	5824	GCAATACATATTGGAGAGAAATGATTAATCTCTGTGTCTGGAGGTGCA	5883	Db	37202	ACGCAATGAGTGCCTGAACTGCACTGCACTGCA	37261
Db	36122	GCAATACATATTGGAGAGAAATGATTAATCTCTGTGTCTGGAGGTGCA	36181	QY	6964	CTGCTCTTGGACCTGACTTCTACCTCTCTGTTACTTGCTCTCGAGATCGAA	7023
QY	5884	GTCAGCTGCTGTTGACTCTTCTACCTACCAACTCTGAACTGTAATCTCC	5943	Db	37262	CTCTCTTGGACCTGTTACTCTCTGTTACTTGCTCTCGAGATCGAA	37321
Db	36182	GTCAGCTGCTGTTGACTCTTCTACCAACTCTGAACTGTAATCTCC	36241	QY	7024	GCTGTRAACTACRAATGAGCCAAATGAGCTTACTGACTTACTTGCTCTCGAG	7083
QY	5944	GATCAGGGAGAGAAAACITGTAAGAGTAATCTCCACCTACCGGGTACATGC	6003	Db	37322	GCTCTTAACTACRAATGAGCCAAATGAGCTTACTGACTTACTTGCTCTCGAG	37381
Db	36242	GATCAGGGAGAGAAAACITGTAAGAGTAATCTCCACCTACCGGGTACATGC	36301	QY	7084	CTGGACCGCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	7143
QY	6004	ACCTCTAGCCCTACAAAGGACTAGATCTCAAACACTGATGAACCCCTCCGTACCC	6063	Db	37382	CTGGACCGCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	37441
Db	36302	ACCTCTAGCCCTACAAAGGACTAGATCTCAAACACTGATGAACCCCTCCGTACCC	36361	QY	7144	GAATCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	7203
QY	6064	ACCTCTAGCCCTACAAAGGACTAGATCTCAAACACTGATGAACCCCTCCGTACCC	6123	Db	37442	GAATCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	37501
Db	36362	ACCTCTAGCCCTACAAAGGACTAGATCTCAAACACTGATGAACCCCTCCGTACCC	36421	QY	7204	SRCRGCCGACCTCCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	7263
QY	6124	ACACCTACTAATCTGTTGATGCTCCCTGACTCTCTGCAATCTTCACTCT	6183	Db	37502	CGTGGCCACCTCCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	37561
Db	36422	ACACCTACTAATCTGTTGATGCTCCCTGACTCTCTGAGCTGAGCTGAGCTGAGCT	36481	QY	7264	AGGACTCTGTTCTAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	7323
QY	6184	GTACCTGAACTGACACTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	6243	Db	37562	AGGACTCTGTTCTAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	37621
Db	36482	GTACCTGAACTGACACTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	36541	QY	7324	ACCTCCCTTAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	7383
QY	6244	CCTCTGTTCCAACTSTGGAAATACCCATACCTCAACCTCACCTGTTAAATTAGC	6303	Db	37622	ACATCCCTTAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	37681
Db	36542	CCTCTGTTCCAACTSTGGAAATACCCATACCTCAACCTCACCTGTTAAATTAGC	36601	QY	7384	ACTAATAGCTTAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	7443
QY	6304	AATACTACATACACCAACTCCAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	6363	Db	37682	ACTAATAGCTTAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	37741
Db	36602	AATACTACATACACCAACTCCAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	36661	QY	7444	AGAGCACAGCAGGAGGAGCAATGATCGGAGTAACTGAGCTGAGCTGAGCTGAGCT	7503
QY	6364	GTCGCTTACCTCAGGAAATTGCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	6423	Db	37742	AGAGCACAGCAGGAGGAGCAATGATCGGAGTAACTGAGCTGAGCTGAGCTGAGCT	37801
Db	36662	GTCGCTTACCTCAGGAAATTGCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	36721	QY	7504	CAACCCCTTGGCTCCCTTATGGAGCTGCTTCACTGAGCTTCACTTCACT	7563

Db	37802	CAACCCCTTGGTCCCCCTGATGGAGCTCTGTTCATGATTCACTCT	37861
Oy	7564	ATTAATCTTGCACTGCR	7582
Db	37862	ATTAATCTTGCACTGCA	37880
RESULT	9		
AC007566/c			
LOCUS	AC007566	149194 bp	DNA linear
DEFINITION	Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.	PRI 01-MAR-2002	
ACCESSION	AC007566		
VERSION	AC007566.2	GI:1118861	
KEYWORDS	HIG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Bukayoya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 149194)		
AUTHORS	Suzlon,J.E. and Waterston,R.		
TITLE	Toward a complete human genome sequence		
JOURNAL	Genome Res.	8 (11), 1097-1108 (1998)	
MEDLINE	99053792		
PUBMED	981074		
REFERENCE	2 (bases 1 to 149194)		
AUTHORS	Du, Z.		
TITLE	The sequence of Homo sapiens BAC clone CTB-10G5		
JOURNAL	Unpublished (2001)		
REFERENCE	3 (bases 1 to 149194)		
AUTHORS	Waterson,N.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-MAY-1999) Genome Sequencing Center, Washington, MO 63108, USA		
REFERENCE	4 (bases 1 to 149194)		
AUTHORS	Waterson,N.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-OCT-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA		
REFERENCE	5 (bases 1 to 149194)		
AUTHORS	Waterson,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	6 (bases 1 to 149194)		
AUTHORS	Waterson,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	7 (bases 1 to 149194)		
AUTHORS	Waterson,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	8 (bases 1 to 149194)		
AUTHORS	Waterson,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA		
COMMENT	On Nov 16, 2000 this sequence version replaced gi:4035815.		
-----	Genome Center		
Center	Washington University Genome Sequencing Center		
Center code	WUGSC		
Web site	http://genome.wustl.edu/gsc		
Contact	sapiens@genome.wustl.edu		
-----	Summary Statistics		
Center Project name	H_RG010605		

NOTICE:	This sequence may not represent the entire insert of this		
FEATURES			
SOURCE			
NEIGHBORING SEQUENCE INFORMATION:			
The clone sequenced to the left is RP5-91H5, 200 base pair overlap. Actual start of this clone is at base position 150532 of CTB-10G5.			
LOCATION/QUALIFIERS			
1..149194			
FEATURES			
SOURCE			
misc_feature			
misc_feature	/organism="Homo sapiens"		
misc_feature	/mol_type="genomic DNA"		
misc_feature	/db_xref="Taxon:9606"		
misc_feature	/chromosome="7"		
misc_feature	/map="q21-7q22"		
misc_feature	/clone="CTB-10G5"		
repeat_region	/clone_id="CTB-978SK-B"		
repeat_region	1..149194		
repeat_region	/rpt_family="L1"		
repeat_region	2248..2388		
misc_feature	/notes="match to EST BG752883 (NID:914063536)"		
misc_feature	2248..2387		
misc_feature	/note="match to EST AW579261 (NID:97254310)"		
misc_feature	2248..2387		
misc_feature	/note="match to EST BG766882 (NID:914077535)"		
misc_feature	2248..2387		
misc_feature	/note="match to EST BT160365 (NID:914620366)"		
misc_feature	2248..2374		
misc_feature	/note="similar to Homo sapiens EST BP758865 (NID:91210665)"		
misc_feature	2248..2287		
repeat_region	/note="match to EST AV686676 (NID:91288539)"		
repeat_region	2296..3066		
repeat_region	/rpt_family="MaLR"		
repeat_region	3108..3392		
misc_feature	/rpt_family="Alu"		
misc_feature	3540..3628		

misc_feature	/note="match to EST AA425526 (NID:g2106267) zw48b03.rl"	Db	90991 CCATTGACCTCAGACACTAAGAAACGACTATTCCTCGAGTCCGCTG 90932
misc_feature	/note="similar to Mus musculus EST BB253526 (NID:g994622)"	Qy	726 GCACTCTGAGGGAGATAATTAAACCACTCTACAGTAGACYCTTTGAGA 785
misc_feature	3542. .3628 /note="match to EST AA149693 (NID:g1720635) znn9d06.rl"	Db	90931 GCACCTCTGAGGAGATAATTAAACCACTCTACAGTAGACYCTTTGAGA 90872
misc_feature	3542. .3628 /note="match to EST AU123510 (NID:g10948226)"	Qy	786 AASGCAGAATGGAGTGCCATAAGAACACTTCTTTCTTAAGAGACACTCAC 845
misc_feature	3542. .3628 /note="match to EST BG766882 (NID:g1407735)"	Db	90871 AAGGCAGAATGGAGTGCCATAAGAACACTTCTTTCTTAAGAGACACTCAC 90812
misc_feature	3542. .3628 /note="match to EST BB272554 (NID:g146913)"	Qy	845 ATTTATTTAAAGTGATTTAGCCCTACAGGAACCTTCAGGTACCTCCATC 905
misc_feature	3542. .3628 /note="match to EST BB299708 (NID:g9183456)"	Db	90811 ATTTATTTAAAGTGATTTAGCCCTACAGGAACCTTCAGGTACCTCCATC 90752
misc_feature	3714. .3785 /note="match to EST BG260659 (NID:g12770475)"	Qy	906 CCAGCATCCCAGCTCTTCCAMYATAAGGACCCCTCAACCAAATGGCAA 965
misc_feature	3717. .3785 /note="similar to Mus musculus EST BB994936 (NID:g1067674)"	Qy	90751 CCAGCATCCCAGCTCTTCCAACTTAAGGACCCCTCAACCAAATGGCAA 90692
misc_feature	3719. .3785 /note="match to EST AA149693 (NID:g1720635) znn9d06.rl"	Db	90631 CCCTT-CGAGTAGGAGGAGATCGGCCGCCAGCAGGTGATGTCGTTTC 90573
misc_feature	3719. .3785 /note="match to EST AA425526 (NID:g2106267) zw48b03.rl"	Qy	1096 TCCCGAGACTTAAGGAAATAAACAGACTTAGTAATTCATAAAVCTGTCG 1145
misc_feature	3719. .3785 /note="match to EST AU123510 (NID:g10948226)"	Db	90572 TCTCCAGACTTAAGGAAATAAACAGACTTAGTAATTCATAAAVCTGTCG 90513
misc_feature	3719. .3785 /note="match to EST BB272564 (NID:g146913)"	Qy	1146 CTATATGTTAGGAAATAAACAGACTTAGTAATTCATAAAVCTGTCG 1205
misc_feature	3719. .3785 /note="match to EST BB299708 (NID:g183456)"	Db	90512 CTATATGTTAGGAAATAAACAGACTTAGTAATTCATAAAVCTGTCG 90454
misc_feature	3719. .3785 /note="match to EST BG766882 (NID:g1407735)"	Qy	1206 TGTACTCTTAATCAGAACCTAACGAACTTCTCGTCACTGAGTTCGCT 1265
misc_feature	3719. .3785 /note="match to EST BI160365 (NID:g14620366)"	Db	90453 TGTACTCTTAATCAGAACCTAACGAACTTCTCGTCACTGAGTTCGCT 90394
misc_feature	3719. .3785 /note="similar to Mus musculus EST BB253526 (NID:g98946272)"	Qy	1266 GAGRTGTTGGGATCTGGTATCTCGTCAGTCAGTCAATGGTGTGACACAGGG 1325
misc_feature	3860. .3862 /note="match to EST AV686676 (NID:g10288539)"	Db	90393 GAGGTGTTGGGATCTGGTATCTCGTCAGTCAGTCAATGGTGTGACACAGGG 90337
repeat_region	4730. .4827 /note="family="L2"	Qy	1326 AAAGANAAATGATCCCAAGGCGAGGARGCGTTCAGCTTACACCTCATGGGGAC 1385
repeat_region	4834. .4925 /note="family="L1"	Db	90336 AAAGAGAAATGATCCCAAGGCGAGGARGCGTTCAGCTTACACCTCATGGGGAC 90280
misc_feature	5715. .5860 /note="match to EST AA425526 (NID:g2106267) zw48b03.rl"	Qy	1386 ACAGAAATCAGAACCTACGGGAAATTGCTCGAGAATTGCTAACTTGTGCTG 1445
misc_feature	5715. .5860 /note="match to EST AU123510 (NID:g10948226)"	Db	90279 ACACAGATCAGACAT-GGAGATGGTGTGAGGATTTGCTGAGA 90221
misc_feature	5715. .5860 /note="match to EST BB272564 (NID:g146913)"	Qy	1446 GGACTAAGGAAGACTTGGAAARTCTAYGATTACTCAATGATGCCACAAATACA 1505
misc_feature	5715. .5860 /note="match to EST BB299708 (NID:g9183456)"	Db	90220 GGACTAAGGAAGACTTGGAAARTCTAYGATTACTCAATGATGCCACAAATACA 90162
misc_feature	5715. .5860 /note="match to EST BG260659 (NID:g12770475)"	Qy	1506 GGGAAAGGAGGAAATCTACTGCCTTCTGGAGAGACTAAGGAGGACATGGAGG 1565
misc_feature	5715. .5860 /note="match to EST BG766882 (NID:g1407735)"	Db	90161 -GGAAGGAAAGAACCTACTGCCTTCTGGAGAGACTAAGGAGGACATGGAGG 90103
misc_feature	5715. .5860 /note="match to EST BI160365 (NID:g14620366)"	Qy	1566 GTCGCTCTGTGACCTGACTCTTCAGGAGCAACTATCTTAAGGAACTTACA 1625
misc_feature	5715. .5860 /note="match to EST BI160365 (NID:g14620366)"	Db	90102 GTCGCTCTGTGACCTGACTCTTCAGGAGCAACTATCTTAAGGAACTTACA 90043
misc_feature	5715. .5860 /note="match to EST BI160365 (NID:g14620366)"	Qy	1626 CTCAGTGTGCTGACGAGATTAGAAACTCAAGTCTCGTAGGCCCGAGCAA 1684
Query Match	85.0%; Score 6448; DB 9; Length 149194; Best Local Similarity 95.7%; Pred. No. 0; Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;	Db	90042 CTCAGTGTGCTGACGAGATTAGAAACTCAAGTCTCGTAGGCCCGAGCAA 89983
Qy	606 GATGGAAACGTCGGAGAACAAAGGCCCTAACAGTATCTGGAGATGGGA 665	Qy	1685 ACTTAGAACCTTGTGACCTGCTGCTTCTGGAGAGACTGAGGAGC 1744
Db	91051 GATGGAAACGTCGGAGAACAAAGGCCCTAACAGTATCTGGAGATGGGA 90992	Db	89982 ACTTAGAACCTTGTGACCTGCTGCTTCTGGAGAGACTGAGGAGC 89923
Qy	666 MCATTGACCCAGACACTAAGAAGAACGACTTATTCCTCTGAGGCCGCTG 725	Qy	1745 AGCGGAGACAGAACGGATTAAAGGCCCTTAGCATGACCTCG 1804

QY	99922	AGGGCGAACAGGACAAACGGGATTAAAAAAGGGCACCCTTACGGCATTC	89853
QY	1805	CANGTGACTTGGGGCTGGAAAGGAAACTGGCCAATGTGAATGCCATTAGG	1864
Db	89862	CAAGTGGACTTTGGCTGGAAAGGAAACCTGGGCAATTGAACTGCTTATAGG	89803
QY	1865	GCTGTGTCAGGGGCTCACAGACATTAAAGATGTGCAAGTAGAGTAG	1924
Db	89802	GCTGTGTCAGGGGCTCACAGACATTAAAGATGTGCAAGTAGAGTAG	89743
QY	1925	CCCCCCTTGTCTGCCCTTATTCAGGGATCACTGAGGCCACTGCCCAGG	1984
Db	89742	CCCCCCTTGTCTGCCCTTATTCAGGGATCACTGAGGCCACTGCCCAGG	89683
QY	1985	GGACAAGCTTGTACTCAGAACCACTAACGATGATGCCAGCAGGACTGAGG	2044
Db	89682	GGACAAGCTTGTACTCAGAACCACTAACGATGATGCCAGCAGGACTGAGG	89623
QY	2045	TGCTGGGCAGGCCATCCATGCCATACCCATACCCATACCCATACCCATAC	2104
Db	89622	TGCTGGGCAGGCCATCCATGCCATACCCATACCCATACCCATACCCATAC	89563
QY	2105	ATTCAGGSCAGGAGGTRCTCTGGACACTGGGGCTCTTAGCTTACCTTC	2164
Db	89562	ATTCAGGSCAGGAGGTRCTCTGGACACTGGGGCTCTTAGCTTACCTTC	89504
QY	2165	TGCCCCACAACCTCCAGCTGTCACTATCTGGGGCTCCNTAGAGGGCA	2224
Db	89503	TGCCCCACAACCTCCAGCTGTCACTATCTGGGGCTCCNTAGAGGGCA	89446
QY	2225	GTCACTAGTACTTYTCCAGCCATACTGTTAGTACTGGGAGCTTATCTTCAC	2284
Db	89445	GTCACTAGTACTTYTCCAGCCATACTGTTAGTACTGGGAGCTTATCTTCAC	89388
QY	2285	ATGTTTCAATTATGTCTGAACGCCACTACCTGTTAGGGAGACATCTAGCAA	2344
Db	89387	ATGTTTCAATTATGTCTGAACGCCACTACCTGTTAGGGAGACATCTAGCAA	89328
QY	2345	AACAGGGCCATTATACCTGACATAGGAGAGAACCCCTTGTGTNCCTG	2404
Db	89327	AACAGGGCCATTATACCTGACATAGGAGAGAACCCCTTGTGTNCCTG	89269
QY	2405	CTTGAGGAAGGATAATTCTGAGCTGGCAAGAAGGACATATGGAGGCCAA	2464
Db	89268	CTTGAGGAAGGATAATTCTGAGCTGGCAAGAAGGACATATGGAGGCCAA	89210
QY	2465	GAATGCCCTCTGTCAGTTAACTTAAGGATTCACTTCTTCCCTACCAAGGA	2524
Db	89209	GAATGCCCTCTGTCAGTTAACTTAAGGATTCACTTCTTCCCTACCAAGGA	89150
QY	2525	GTACCCCTGAGCCAAAGGCCAACAGGATCTTGTAGGACTTAAAGC	2584
Db	89149	GTACCCCTGAGCCAAAGGCCAACAGGACTTAAAGGAACTTAAAGC	89090
QY	2585	CCAGGCTAGTAAMACCATGCACTCCCTGCGATAATTCCGTAGTGGATTGAGG	2644
Db	89089	CCAGGCTAGTAAMACCATGCACTCCCTGCGATAATTCCGTAGTGGATTGAGG	89037
QY	2645	CACAGAACCCAGTGACAGTGACAGTGACAGTGACAGTGACAGTGACAG	2704
Db	89036	TACAGAACCCAGTGACAGTGACAGTGACAGTGACAGTGACAGTGACAG	88979
QY	2705	CGTCTCTTATACCCAGTGCTACCGCTTATCTGTGMYTCCAATCCAGA	2764
Db	88978	TGTCTCTCTATGCACTGAGCTTACCTGAGATCTGAGATCTGAGATCTGAG	88919
QY	2765	GGAAACAGAGGTGGTTACASTCTCTGGAGCTTACAGATGCTCTTGCA	2824
Db	88918	GGAAACAGAGGTGGTTACASTCTCTGGAGCTTACAGATGCTCTTGCA	88859
QY	2825	TCCTRACTCTCAATCTGTGTGCTTGAGATCTGAGATCTGAGATCTGAG	2884
Db	88858	TCCTRACTCTCAATCTGTGTGCTTGAGATCTGAGATCTGAGATCTGAG	88799
QY	2885	CTGAGCTTCTTACCCAAAGGGTCAAGGATAGYCCOATTATGGCCGATTC	2944
Db	88798	CTGAGCTTCTTACCCAAAGGGTCAAGGATAGYCCOATTATGGCCGATTC	88739
QY	2945	CCAGACTTGGYCARITYMTCATCTGGACACTCTGCTCTCGTAKGATGATT	3004
Db	88738	CCAGACTTGGYCARITYMTCATCTGGACACTCTGCTCTCGTAKGATGATT	88681
QY	3005	ACTTTRGCVCCYRITCAGAACCTGTGTCATAGGCCCAAGRCRCTCTMAATT	3064
Db	88680	ACTTTRGCVCCYRITCAGAACCTGTGTCATAGGCCCAAGRCRCTCTMAATT	88621
QY	3065	OCTCGCTACCTGTGGTACATGGTTCCAAACCAAGGTCACCTGCTCACAGCAGT	3124
Db	88620	OCTCGCTACCTGTGGTACATGGTTCCAAACCAAGGTCACCTGCTCACAGCAGT	88561
QY	3125	TAATCTCTAGCTTACATGATCCAGGAACTTACAGGATCCAGGAGCTTAC	3184
Db	88504	TAATCTCTAGCTTACATGATCCAGGAACTTACAGGATCCAGGAGCTTAC	88445
QY	3245	AVAGSITTCGGGAAWATGATCCAGGCTTGTGGCAGATAGCCAGGAT	3304
Db	88444	AVAGSITTCGGGAAWATGATCCAGGCTTGTGGCAGATAGCCAGGAT	88386
QY	3305	CASTANTTAAGGAAACTCTAGAACAGGCCATACCCATTARTAGGATGAA	3364
Db	88385	CASTANTTAAGGAAACTCTAGAACAGGCCATACCCATTARTAGGATGAA	88280
QY	3365	MAGTGCSTTCCAGGCCCTAAAGGCTTAACCCAAAGGCCAGTGTAGYTGCC	3424
Db	88325	MAGTGCSTTCCAGGCCCTAAAGGCTTAACCCAAAGGCCAGTGTAGYTGCC	88226
QY	3425	ACRGGGAAACCTTSTTATATTCAGAACCCATACCTGATAGGGCAGTCATT	3484
Db	88279	ACRGGGAAACCTTSTTATATTCAGAACCCATACCTGATAGGGCAGTCATT	88221
QY	3485	TTCACAGRTCTCAGGAGYGGCTTGCACCTCYRACCTGAGTGAAYTGATG	3544
Db	88220	TTCACAGRTCTCAGGAGYGGCTTGCACCTCYRACCTGAGTGAAYTGATG	88161
QY	3545	TAGTGCCTAAAGGTTCTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTG	3604
Db	88160	TAGTGCCTAAAGGTTCTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTG	88101
QY	3605	CTGAGGAGTAAATATAGGRAGAGACTTACTGTGAGACATCTAGAAGTGA	3664
Db	88100	CTGAGGAGTAAATATAGGRAGAGACTTACTGTGAGACATCTAGAAGTGA	88041
QY	3665	ATYRGCACTACTCTGCTAACAGGAGCTTGTGGCTGTGAGACATCTAGA	3724
Db	88040	ATYRGCACTACTCTGCTAACAGGAGCTTGTGGCTGTGAGACATCTAGA	87981
QY	3725	AGGCTTCTACTCTGCTAACAGGAGCTTGTGGCTGTGAGACATCTAGA	3784
Db	87980	AGGCTTCTACTCTGCTAACAGGAGCTTGTGGCTGTGAGACATCTAGA	87921
QY	3785	YENCATGCTTCAAGATGAGAGATGAGATGAGATGAGATGAGATGAGAT	3844
Db	87720	YENCATGCTTCAAGATGAGAGATGAGATGAGATGAGATGAGATGAGAT	87861
QY	3845	AAACCTATGCCTCTGGAGGACCTTGTGGCTGTGAGCTGCTTACCTTACCCAG	3904
Db	87860	AAACCTATGCCTCTGGAGGACCTTGTGGCTGTGAGCTGCTTACCTTACCCAG	87802
QY	3905	TWTATACTGTGAGACTTCTTGTGAAAGGACTCTGAGAAGYGGGTATGAGT	3954
Db	87801	TWTATACTGTGAGACTTCTTGTGAAAGGACTCTGAGAAGYGGGTATGAGT	87742

Db	85582	AACCTACTAATCTGTTGATGCTCCCTGAACCTGGCCATATGTTCAATCCCT	85523	QY	7264	AGGACTGCTGATTCTAGGCTGATAGAACATCCTAGGGAAAGGTGACC	7323
QY	6184	GTAACCTGAAACATGGAAACACTTCAGCACAGAAATAACCCACCTCCCGTTTAGGA	6243	Db	84442	AGGACTAATCTGATTTCTCTAGGCTGATAGAACATCCCTAGGCTAGCTGGAAAGGTGACC	84383
Db	85522	GTAACTGAAACATGGAAACACTTCAGCACAGAAATAACCCACCTCCCGTTTAGGA	85463	QY	7324	ACATCCACCTTAAACAGGGGCTGAACTGTCACCTGACCTTGAACTGAGGCTC	7383
QY	6244	CCTCTGTTCCAAATSTGAAATAACCAACCTCAACCTCACCTGTTAAATTAGC	6303	Db	84382	ACATCCACCTTAAACAGGGGTTGAACTGACTGTCACCTGACCTGAGGCTC	84323
Db	85462	CCTCTGTTCCAAATCTGGAATAACCAACCTCAACCTCACCTGTTAAATTAGC	85403	QY	7384	ACATTAATGCTAAATGCAAGACACAGGAAAGATAAGCAATCATYATATGCMTC	7443
QY	6304	AATCTACATACACCAACTCCAACTGATCAGGGTTAATCTCCACAGAAATA	6363	Db	84322	ACATTAATGCTAAATGCAAGACACAGGAAAGATAAGCAATCATYATATGCMTC	84263
Db	85402	AATCTACATACACCAACTCCAACTGATCAGGGTTAATCTCCACAGAAATA	85343	QY	7444	AGAGCACAGCAGGGCAATGATCGGATATAACCAAGGTTGAGCCGCAACCG	7503
QY	6364	GTCCTGCCTACCTCTAGGAAATTCTTGTGTTGTTGATGTTGTGAT	6423	Db	84262	AGAGCACAGCAGGGCAATGATCGGATATAACCAAGGTTGAGCCGCAACCG	84203
Db	85342	GTCCTGCCTACCTCTAGGAAATTCTTGTGTTGTTGATGTTGTGAT	85283	QY	7504	CAACCCCTTGGTCCCTCCCTTATGGAGCTGTTCACTTATTCCT	7563
QY	6424	GGCTCTTCAGAACTATGCTCTCTATCTTAGGCCCCATGTCCTACACT	6483	Db	84202	CAACCCCTTGGTCCCTCCCTTGTATGSGAGCTGTTCACTTATTCCT	84143
Db	85282	GGCTCTTCAGAACTATGCTCTCTATCTGTCCTACACT	85223	QY	7564	ATTAATCTGCACTGCR	7582
QY	6484	GAACAGAGTTTACGTTTGTCAATTAAGCCGACAAAGGTTACCCATT	6543	Db	84142	ATTAATCTGCACTGCR	84124
Db	85222	GAACAGAGTTTACGTTTGTCAATTAAGCCGACAAAGGTTACCCATT	85163	RESULT 10			
QY	6544	CCTTTGTTAGGAGCAGGAGTGTAGGTGACTAGGTACTGGATCGGGTATCA	6603	AY01583			
Db	85162	CCTTTGTTAGGAGCAGGAGTGTAGGTACTGGATCGGGTATCA	85103	DEFINITION	AY01583		
QY	6604	ACCTCTACTAGTCTACTCTCAAACATCTCAAGACTAAATGGGAGCTGAAAGGGTC	6663	ACCESSION	AY01583	Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERWE1 locus, allele B, complete sequence.	
Db	85102	ACCTCTACTAGTCTACTCTCAAACATCTCAAGACTAAATGGGAGCTGAAAGGGTC	85043	VERSION	AY01583.1	GI 3754401	
QY	6564	GCCCCACTCCCTGGTACCTTGCAAGATCACTCTCTCTAGGAGCTGTTTTAGG	6723	KEYWORDS			
Db	85042	GCCCCACTCCCTGGTACCTTGCAAGATCACTCTCTCTAGGAGCTGTTTTAGG	84983	SOURCE			
QY	6724	ATATGAGAGCTTAGACTGCTAACCGCTGARAGGGGAACTCTGTTTTAGG	6783	ORGANISM		Homo sapiens (human)	
Db	84982	ATATGAGAGCTTAGACTGCTAACCGCTGARAGGGGAACTCTGTTTTAGG	84923	REFERENCE		Bukay-York, Metazo; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butharia; Primates; Catarhini; Hominidae; Homo.	
QY	6784	GAAGAGATGCTGTTATATGTTATGCAATCGGAATGTCCTGAGAAAGTRAAATT	6843	AUTHORS		Mallet, F., Bouton, O., Prudhomme, S., Chemet, V., Oriol, G., Bonnaffon, B., Lugeon, G., Duret, L. and Mandrand, B.	
Db	84922	GAAGAGATGCTGTTATATGTTATGCAATCGGAATGTCCTGAGAAAGTRAAATT	84863	TITLE		The endogenous retroviral locus ERWE1 is a bona fide gene involved in hominoid placental physiology	
QY	6844	CSAGATCGATAACAGTAAGCAGARGACTGCAACCTGAACTGCACTGGGCTCTC	6903	JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	
Db	84862	CGAGATGCTGTTATATGTTATGCAATCGGAATGTCCTGAGAAAGTRAAATT	84803	PUBMED		1475826	
QY	6904	ACCRATGGTCCCTGGATCTCCCTTTAGACTCTAGAGCTTAAATATGCA	6963	REFERENCE		2 (bases 1 to 1022)	
Db	84802	ACCCRATGGTCCCTGGATCTCCCTTTAGACTCTAGAGCTTAAATATGCA	84743	AUTHORS		Mallet, F., Bouton, O. and Oriol, G.	
QY	6964	CTCCCTCTTGGACCCGTATTTACTCTGTTACTGTCCTTCAGATGCA	7023	TITLE		Direct Submission	
Db	84742	CTCCCTCTTGGACCCGTATTTACTCTGTTACTGTCCTTCAGATGCA	84683	JOURNAL		Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS Biomerieux, Ecole Normale Supérieure de Lyon - 46 alle	
QY	7024	GCTGTTAACTACATAATGGGCCAGATCAGTCAGTCAGTAAGCTAACGGACCC	7033	FEATURES		d'Italie, Lyon 69364 cedex 07, France	
Db	84682	GCTGTTAACTACATAATGGGCCAGATCAGTCAGTCAGTAAGCTAACGGACCC	84623	SOURCE		Location/Qualifiers	
QY	7084	CTGGACCGGGCTGTTAGGCCAGATCAGTCAGTCAGTAAGCTAACGGACCC	7143		1. 1022		
Db	84622	CTGGACCGGGCTGTTAGGCCAGATCAGTCAGTCAGTAAGCTAACGGACCC	84563		/organism="Homo sapiens"		
QY	7144	GAAATCTCAGCTGCAACCTCTACGCCCCAAATCGAGAGAGCTAGGGGT	7203		/mol_type="genomic DNA"		
Db	84562	GAAATCTCAGCTGCAACCTCTACGCCCCAAATCGAGAGAGCTAGGGGT	84503		/isolate="21"		
QY	7204	SGTGGCCAACTCCCAACAGACTCTAGGTTCTGTTAGATGGGGAACTGAGAGAC	7263		/isolation_source="PBMC of individual 21"		
Db	84502	SGTGGCCAACTCCCAACAGACTCTAGGTTCTGTTAGATGGGGAACTGAGAGAC	84443		/chromosome="7"		
				/map="q21-q22"			
				/sex="Female"			
				/cell_type="PBMC"			
				/notes="isolated by PCR; ERWE1 locus, allele B			
				/endogenous_virus: HERV-W"			
				1. .780			
				/join(248..811,2887..9813)			
				/notes="putative mRNA transcript 1; gag-pol pseudogene"			
				/join(48..811,7585..9813)			
				/product="envelope glycoprotein"			
				/notes="putative mRNA transcript 2"			
				/join(48..811,9248..9813)			
				/notes="putative mRNA transcript 3; no coding region"			
				/join(7812..9428)			

```

/note="syncyrin"
/codon_start=1
/protein="envelope glycoprotein"
/protein_id="RAB6162_1"
/db_xref="GI|3754402"
/translation="MALPYHIFPFVILPSPFTTAPPCRMWSSPQEFWRMOP
GNIDAPSRLSKGTPPTPAHTIMPYHSATLCHANTWPMINSCGGGV
TWCWTFQIGNSDGGVQDORAKHEKAEVLSQTRVAGTSBPKGIDLSKHATLR
HTFLVSLNTTTLGILHEVAONTNCHCPLNFRPVISPEBWNPNSTINTSV
LVPGLVLSNLLEITPTTGTGKVSQDPMQIWTQJUCLSGIFFVGTSB
YRCLGSSSESMCLSLSPMPMTYEODLYSYVLSKPKNRVPLPVIAGYTGALG
TGGGIFTMSTOFYKLSLNGEMLNERVADSLVTDOLNSLAARVLLNRRALDLTA
RGTCLEFGEBCCYVYUNQSVIYTKVEIRDRIQRAELRNGPGWGLSQMPWLP
FLAGPLAATILLIIFGPCIFNLNNLFVASSRIEAVKLOMPKOMSKTKIYRRPLURPASP
RSVDNDIKGTPREISAQQLPRLPNSAGSS"

```

ORIGIN LTR

Query	Match	Score	Length
Best Local Similarity	85.0%	6446.4;	10222;
Matches	6680;	Pred. No.	0;
Conservative	95.7%;	Mismatches	51;
		Indels	53;
		Gaps	27;
Qy	606	GATCGGAAGCTTCCCGAGAACAGACAAAAGGCCCTAAGCGTATTCTGGAA	665
Db	2896	GATCGGAAGCTTCCCGAGAACAGACAAAAGGCCCTAAGCGTATTCTGGAA	2945
Qy	666	MCATTTGACCTCTGAGACACTTAAGAAGAACGACTTATTTCTCGAGARATGGAA	725
Db	2946	CCATTGACCTCTGAGACACTTAAGAAGAACGACTTATTTCTCGAGARATGGAA	3005
Qy	726	GCACCTCGAGGGAGATAATTACACCATCTTACGACTGAGCTCTTGTAGA	785
Db	3006	GCACCTCTGAGGGAGATAATTACACCATCTTACGACTGAGCTCTTGTAGA	3065
Qy	786	AAGGCAATGAGGTGAAGTGCATAGAACACTTCTTCTTCTTAAGAGAACACTAC	845
Db	3066	AAGGCAATGAGGTGAAGTGCATAGAACACTTCTTCTTCTTAAGAGAACACTAC	905
Qy	846	AATTTATGTTAAAGTGTGATTATGCCATAGAACACTTCTTCTTCTTAAGAGAACACTAC	3125
Db	3126	AATTTATGTTAAAGTGTGATTATGCCATAGAACACTTCTTCTTCTTAAGAGAACACTAC	3185
Qy	906	CCAGCATCCCAGACTCTTCCCAMYTATAAGACCCCTCACCCAATGTCGA	965
Db	3186	CCAGCATCCCAGACTCTTCCCRACTATTAAGACCCCTCACCCAATGTCGA	3245
Qy	966	AAGGAGATAGACAAGGTTAACAGTGCACCCAAAGAGGATCTGGCCATATCCCCATTATGA	1025
Db	3246	AAGGAGATAGACAAGGTTAACAGTGCACCCAAAGAGGATCTGGCCATATCCCCATTATGA	3305
Qy	1026	CCCTCTCAGGAGCTGGAGGAAGAGAACGTTGGCCACAGAGTCCATATCCCAATTATGA	1085
Db	3306	CCCTCTCAGGAGCTGGAGGAAGAGAACGTTGGCCACAGAGTCCATATCCCAATTATGA	3364
Qy	1086	TCTCCAGACTTAAGCAATAAACAGACTTAGTAAATTCAGATAVCTGATG	1145
Db	3365	TCTCCAGACTTAAGCAATAAACAGACTTAGTAAATTCAGATAVCTGATG	3424
Qy	1146	CTATATTGTTGTTACAGGTTAACCCCAATAGAGAGAGTCTTGATGACATGGAGATATA	1205
Db	3425	CTATATTGTTGTTACAGGTTAACCCCAATAGAGAGAGTCTTGATGACATGGAGATATA	3483
Qy	1206	TGTCACTGCTTAATCAGACATACCCCAATAGAGAGAGTCTTGATGACATGGAGATATA	1265
Db	3484	TGTCACTGCTTAATCAGACATACCCCAATAGAGAGAGTCTTGATGACATGGAGATATA	3543
Qy	1266	GAGRGTGTTGGCACTCTGATCTCAGTCACTGATGATGAGAGGAG	1325
Db	3544	GAGAGTGTGGCACTCTGATCTCAGTCACTGATGAGAGGAG	3600
Qy	1326	AAAGANAAATGATTCCTCAAGGGCAGCAGCAGTCAGCTTCCAGTCAGCT	1385
Db	3601	AAAGAGAGATGATTCCTCAAGGGCAGCAGCAGTCAGCTTCCAGTCAGCT	3657

Qy	4728	GAATGCCGCTCTTCAGTAACTTCAAGTTCAACTTCCCTACCAAGCA	2524	Db	5777	TATGGCAGAGGTGACTCTTGTTACGGTAGTGGCGAGTCAGCTTGTAGT	5833
Qy	4788	GTAACCCTCAGACCAGGCCAACAGGATTCAGGTTAAGCTAAAGC	2584	Db	3605	CYGAAGAGTTAAATACTGGGRAGAGCTACTGTGGACATCTCTGAKGTA	3664
Qy	4788	GTAACCCTCAGACCAGGCCAACAGGATTCAGGTTAAGCTAAAGC	2584	Db	5837	CTGAAGGAGTTAAATACTGGAGAGATCTACTGTGGACATCTGATGTA	5896
Qy	2645	CAAGGAACCCAGGGAGCTGAGGTGA-GGTAGTGCAGAATGAGG	4847	Db	3665	AAGCAACTCTGTAAGGAGACTGTGCTAGACACGCTTACTAARTC	3724
Qy	4901	TACAGAACCCACAGAGTGA-GGTAGTGCAGAATGAGG	4900	Db	5997	AEGCAACTCTGTAAGGAGACTGTGCTAGACACGCTTACTAARTC	5956
Qy	2705	CGTTGCTTTATCCAGCTGAACTTACCGCTTATCTGAGATCTGAG	2704	Db	3725	AEGCAACTCTGTAAGGAGACTGTGCTAGACACGCTTACTAARTC	3784
Qy	4959	TGTGTTCTCTATAGCCAGCTGACTAGCCCTATACCTGCTTCCAA	5018	Db	5957	AEGCAACTCTGTAAGGAGACTGTGCTAGACACGCTTACTAARTC	6116
Qy	2765	GGAGCAGACTGTTACSTCCGACCTMAGGATCCTTCGATCCTGACA	2824	Db	6017	AEGCAACTCTGTAAGGAGACTGTGCTAGACACGCTTACTAARTC	6076
Db	5019	GGAGCAGACTGTTACCTGAGCTTACCCGCTTCTGATCCTGACA	5078	Db	3845	AACCTATGCCCCTGAGGGCTTAGAGTCTGAGCTGACTGCTTACCC	3904
Qy	2825	TCTGACTCTCAATTCTGTTGCTTGCCTTGCAGATCTCAACCC	2884	Db	6077	AACCTATGCCCCTGAGGGCTTAGAGTCTGAGCTGACTGCTTACCC	6135
Db	5079	TCTGACTCTCAATTCTGTTGCTTGCCTTGCAGATCTCAACCC	5138	Db	3905	TGTATACTGTTGAGAAGTCTCTGAGAATAGTAACTCTAACAA	3964
Qy	2885	CTCGACTTACCCAGGGTAGGATGAGATCTCAACCC	2944	Db	6116	TGTATACTGTTGAGAAGTCTCTGAGAATAGTAACTCTAACAA	6195
Db	5139	CTCGACTTACCCAGGGTAGGATGAGATCTCAACCC	5198	Db	3985	CTCGACTTACCCAGGGTAGGATGAGATCTCAACCC	3984
Qy	2245	CCAGACTGAGTCARTWTCATACCTGACACTCTTCTGAGTAKG	3004	Db	6196	GTCTGTTGAGAATGAGATCTCAACCC	6255
Db	5199	CCAGACTGAGCTTACCTCATCTGAGACA-CTTGCTCTGGATG	5256	Db	4025	CTCGACTTACCCAGGGTAGGATGAGATCTCAACCC	4084
Qy	3005	ACITTRGCGCYCURTTCAAGAACCTTGGCCATCAAGCACCA	3064	Db	6256	CTCGACTTACCCAGGGTAGGATGAGATCTCAACCC	6315
Db	5257	ACTTTGGCCCATTCAGAACCTTGGCCATCAAGCACCA	5316	Db	4085	TATACAGACTCTTACCTGAGATCTGAGCTTACAGCAGT	4024
Qy	3065	CCTCGCYACCTGCTGCTACAWGGTTCCAAACCSRARGCT	3124	Db	6316	TATACAGACTCTTACCTGAGATCTGAGCTTACAGCAGT	6375
Db	5317	ACTCTGCTACTCTGCTGCTACATGGTTCCAAACCCAAGGTCA	5376	Db	4145	AGGAAATTCTACTCTGAGGAGACCTCTACAGCTACAGCAGT	4144
Qy	3125	TAATACTTAGGCTAATATTCCAAAGGCACARGCCCTCACTGAG	3184	Db	6376	AGGAAATTCTACTCTGAGGAGACCTCTACAGCTACAGCAGT	6435
Db	5377	---TACTTAGGCTAAATTATCCAAAGGCACCGGGCCTCTAGGAA	5432	Db	6496	TATGGCTGAGGAGGAACTTCTAGAGGTTGAGCTTACAGCAGT	6495
Qy	3185	GCCTATACCTGCTTACCTCATCUCAAACCTTAAGGACTAAGG	3144	Db	6436	TATGGCTGAGGAGGAACTTCTAGAGGTTGAGCTTACAGCAGT	6495
Qy	5413	GCCTATACCTGCTTACCTCATCUCAAACCTTAAGGACTAAGG	5492	Db	4265	GAAGAAGAAAGGAAATASAGRGAAAYTGCCTACAGCATATG	4324
Db	5413	GCCTATACCTGCTTACCTCATCUCAAACCTTAAGGACTAAGG	5493	Db	6496	GAAGAAGAAAGGAGGAACTTCTAGAGGTTGAGCTTACAGCAGT	6555
Qy	3245	AACTGGTTCTGCGCAAATGATGCCAGTWGCTGCTGCTGCT	3304	Db	4325	AGGAGGACCTCTAGAGGAACTTCTAGAGGTTGAGCTTACAGCAGT	4384
Qy	5493	AACTGGTTCTGCGCAAATGATGCCAGTWGCTGCTGCTGCT	5551	Db	6556	GAAGAAGAAAGGAGGAACTTCTAGAGGTTGAGCTTACAGCAGT	6615
Db	3305	CASTRATTAGGAAACTCAGAACGCAATTCTAGATGGAYAMCTG	3364	Db	4385	GGGAAACAGGCCACTACTCGAGGAGAACAGATGGGAACTTCAG	4443
Qy	5552	CACTATTAGGAAACTCAGAACGCAATTCTAGATGGAYAMCTG	5611	Db	6616	GGGAAACAGGCCACTACTCGAGGAGAACAGATGGGAACTTCAG	6675
Qy	3365	AAGTGGCTTCAGGCCACTACTCGAGGAGAACAGATGGGAACTTC	3424	Db	4444	TTCCTCCCTCGGACGCTTACCTCATCAAACCTTCTACTGAGCT	4503
Db	5612	AAGTGGCTTCAGGCCACTACTCGAGGAGAACAGATGGGAACTTC	5657	Db	6676	TTCCTCCCTCGGACGCTTACCTCATCAAACCTTCTACTGAGCT	6735
Db	3485	AACRGGGCAACTTSTYATRTCAAGAAACAGRAVAGCTCTGGAG	3484	Db	4504	CTATGGAATTAACCTCATCAAACCTTCTACTGAGCT	4563
Qy	5658	AACGGGGCAACTTCTGCTACAG-AAAAGAGGATAGCTCTAGGTC	5716	Db	6736	CTATGGAATTAACCTCATCAAACCTTCTACTGAGCT	6795
Db	5717	TTACACAGATCCGAGGGATAGCTCTGTCACCTGCTACAGGA	5776	Db	4564	TATAGGGCAACTTACTGAGGAGGCTTCAACTATGAGC	4623
Qy	3545	TAGGGCAAAAGGGTGTGRCYCATGTTAAGGGTGTGGGGACT	3604	Db	6796	TAGGGCAAAAGGGTGTGRCYCATGTTAAGGGTGTGGGGACT	6855
Qy				Db	4624	CAGGGCTGAACTGCTGCAACTATGAGGAACTTCAGCTCCTCAG	4683

	AUTHORS	Mallet, F., Bouton, O. and Oriol, G.
	TITLE	Direct Submission
	JOURNAL	Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
	FEATURES	CNRS-Biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France
source		Location/Qualifiers 1..10222
QY	AGCCATGGATGCCCTTGTAGGACCTTAGCAACTATATGCTA	9134
Db	CTCCTTGTGAGGCTGTATACTTCCTCTTGTAGGACCTTAGCAACTATATGCTA	6963
9135	AGCCATGGATGCCCTTGTAGGACCTTAGCAACTATATGCTA	9194
QY	GCTGTTAACTACAATGGGCCAGATGCTCAGACTAATATGCTA	7023
Db	GCTGTTAACTACAATGGGCCAGATGCTCAGACTAATATGCTA	9254
QY	CTGGACGGCTGTGAGCCAGATGCTGTTAATGCTCAGACTAATATGCTA	9254
Db	GTAATTCAGTGACGACCTACTACGGCCAACTCGAGGAGCAGTAGGGGT	7083
9255	GTAATTCAGTGACGACCTACTACGGCCAACTCGAGGAGCAGTAGGGGT	9314
QY	CTGGACGGCTGTGAGCCAGATGCTCAGACTAATATGCTA	7143
Db	GAACTCTGACGCTGAGCCAGATGCTGTTAATGCTCAGACTAATATGCTA	9374
9315	GAACTCTGACGCTGAGCCAGATGCTGTTAATGCTCAGACTAATATGCTA	9374
QY	GAAATTCAGTGACGACCTACTACGGCCAACTCGAGGAGCAGTAGGGGT	7203
Db	GAAATTCAGTGACGACCTACTACGGCCAACTCGAGGAGCAGTAGGGGT	9434
9375	GAAATTCAGTGACGACCTACTACGGCCAACTCGAGGAGCAGTAGGGGT	9434
QY	SGTCGGCAACCTCCCCAACGCACTTAGTCAGTCACCTGACCATAGAGCT	7263
Db	SGTCGGCAACCTCCCCAACGCACTTAGTCAGTCACCTGACCATAGAGCT	9494
9435	SGTCGGCAACCTCCCCAACGCACTTAGTCAGTCACCTGACCATAGAGCT	9494
QY	7264 AGGACTAGTGATTCTCTGGCTGATGAGCTTAACTGAGCTAGGAGCT	7323
Db	7264 AGGACTAGTGATTCTCTGGCTGATGAGCTTAACTGAGCTAGGAGCT	9554
9495	AGGACTAGTGATTCTCTGGCTGATGAGCTTAACTGAGCTAGGAGCT	9554
QY	7204 ACATCACCTTAACACGGGCTGACCTAGTCACCTGACCATAGAGCT	7383
Db	7204 ACATCACCTTAACACGGGCTGACCTAGTCACCTGACCATAGAGCT	9614
9555	ACATCACCTTAACACGGGCTGACCTAGTCACCTGACCATAGAGCT	9614
QY	7384 ACTAAATGTTAATTAGGCAAGACGGGAAAGAAATAGCCATATTTGCTG	7443
Db	9615 ACTAAATGTTAATTAGGCAAGACGGGAAAGAAATAGCCATATTTGCTG	9674
QY	7444 AGAGCACAGAGGAGGAGCATGATGGGATAAACCCAGTYTGACCGGAGCG	7503
Db	9675 AGAGCACAGAGGAGGAGCATGATGGGATAAACCCAGTYTGACCGGAGCG	9734
QY	7504 CAACCCCTTGGGCCCTCCTTGTATGGAGCTCTTTCTGATTGCTT	7563
Db	9735 CAACCCCTTGGGCCCTCCTTGTATGGAGCTCTTTCTGATTGCTT	9794
QY	7564 ATTAATCTGCARCTGCR	7582
Db	9795 ATTAATCTGCACCTGCA	9813
RESULT 11	ORIGIN	LTR
AY101584	AY101584	10222 bp
LOCUS	Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERWE1 locus, allele A, complete sequence.	DNA
DEFINITION	Linear	PRI 11-FEB-2004
ACCESSION	AT101584	
VERSION	AY101584.1	GI:37544403
KEYWORDS	.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
BEST LOCAL SIMILARITY	84.9 %;	Score 6440; DB 9; Length 10222;
QUERY MATCHES	6676;	Conservative 155; Mismatches 55; Indels 53; Gaps 27;
QY	606 GATGGGAAGTCCCGAGAACAAAGGCCCTAAGCTGATTCGAAATTGGGA	665
Db	2886 GATGGGAAGTCCCGAGAACAAAGGCCCTAAGCTGATTCGAAATTGGGA	2945
QY	666 MCATTTGACCTCTGAGACTAAGAACGACTATATTCTCTGAGTCGGCGCTG	725
Db	2946 CCAATTGACCTCTGAGACTAAGAACGACTATATTCTCTGAGTCGGCGCTG	3005
QY	726 GCACTCTGAGGGAGTAAATTAACTACCATCTACGCTAGAGCAGCTTGTAGAA	785
Db	3006 GCACCTCTGAGGGAGTAAATTAACTACCATCTACGCTAGAGCAGCTTGTAGAA	3065
QY	786 AAGGAAATGAGTGAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC	845
Db	3066 AAGGAAATGAGTGAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC	3125
QY	846 AATTATGAAAAGTGAGTGTGATTGCTCAGGAGCTCAGAGTCAGTCAGTC	905
Db	3126 AATTATGAAAAGTGAGTGTGATTGCTCAGGAGCTCAGAGTCAGTCAGTC	3185
JOURNAL	In hominid placental physiology	
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	
REFERENCE	1 (bases 1 to 10222) Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo. Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnau, B., Lucotte, G., Duret, L. and Mandrand, B. The endogenous retroviral locus ERWE1 is a bona fide gene involved in hominid placental physiology	
REFERENCE	2 (bases 1 to 10222)	
Db	3186 CCAGCTCCCGACTCTCCCAACTAATAGGCCCCCTGACCAAATGGCTCAA	3245

QY	966	AAGGAGATACACAAGGGTAAACAGTGAAACCAAGAGTCCTAATTCCCCAATATGAA	1025	QY	2045	TGCCTGGCCAAAGGCCATCCCATGCCTCACCTCACAGGSCCTTGGCTATGCTGACC	2104	Db	4315	TGCCTGGCCAAAGGCCATCCCATGCCTCACCTCACAGGSCCTTGGCTATGCTGACC	4374				
Db	3246	AAGGAGATACACAAGGGTAAACAGTGAAACCAAGAGTCCTAATTCCCCAATATGAA	3305	QY	2105	ATTCAGGSCCAGGAAAGGTGTCCTCTSGAACCTGTCATGCTGCTTACTTC	2164	Db	4375	ATTCAGGSCCAGGAAAGGTGTCCTCTSGAACCTGTCATGCTGCTTACTTC	4333				
QY	1026	CCCTTCCCAGCAGTGGGAGAGAATGGGCCAGCCAGACTGCTGCTGCTTACTTC	1085	QY	2165	TGRCGCCAGAACCTGTCCTCGATCTGTCATCTGAGGGCTTNTAGGGCA	2224	Db	4375	ATTCAGGSCCAGGAAAGGTGTCCTCTSGAACCTGTCATGCTGCTTACTTC	4333				
Db	3306	CCCTTCCCAGCAGTGGGAGAGAATGGGCCAGCCAGACTGCTGCTGCTTACTTC	3364	QY	3425	CTATATTGATGTTTACAGGGTTAGGAACTCTTGATGCTGACATGGAGATA	3483	QY	4434	TGRCGCCAGAACCTGTCCTCGATCTGTCATCTGAGGGCTTNTAGGGCA	4391				
QY	1086	TCTCCAGACTTAAGCAAACTAAACACAGACTTAGTAATTCTCAGATAACCTGATG	1145	Db	3365	TCTCCAGACTTAAGCAAACTAAACACAGACTTAGTAATTCTCAGATAACCTGATG	3424	QY	2225	GTCACTGATACCTTCCAGGTTAGGACATTCTTGATCTGACATGGAGATA	2284	Db	4492	GTCACTGATACCTTCCAGGTTAGGACATTCTTGATCTGACATGGAGATA	4549
Db	3425	CTATATTGATGTTTACAGGGTTAGGAACTCTTGATGCTGACATGGAGATA	3483	QY	1206	TGCACTGTTAACTCAGAAGACTACCCAAATGAGAGAATGGCCACCTACTGAGCT	1265	QY	2285	ATGCTTCTTAATATGTTGAAAGCCCACTACTCTTGTAGGAGAGACATCTGAC	2344	Db	4550	ATGCTTCTTAATATGTTGAAAGCCCACTACTCTTGTAGGAGAGACATCTGAC	4609
QY	1266	TGCACTGTTAACTCAGAAGACTACCCAAATGAGAGAATGGCCACCTACTGAGCT	1325	QY	2445	AGCAGGGCCATTATCACCTGAACTTAGGAAGGAGAACGACCCACTACTTGTAGGAGAG	2404	Db	4510	AAGCAGGGCCATTATCACCTGAACTTAGGAAGGAGAACGACCCACTACTTGTAGGAGAG	4668				
Db	3544	TGCACTGTTAACTCAGAAGACTACCCAAATGAGAGAATGGCCACCTACTGAGCT	3600	QY	2405	CTTGAGGAGGAAATTCTCAGTCAGCTGAGACTTGGGACATTTGGGAC	2464	Db	4669	CTTGAGGAGGAAATTCTCAGTCAGCTGAGACTTGGGACATTTGGGAC	4727				
QY	1326	AGAGAGATGATCCCAAGGCCAGCARGCTCCAGCTTASACCTCATTTGGGAC	1385	QY	2465	GAATGCCGCTCTGTCAGTAAGTAACCTAACGATTCCACTCTTCTTACCAAGGCA	2524	Db	4728	GAATGCCGCTCTGTCAGTAAGTAACCTAACGATTCCACTCTTCTTACCAAGGCA	4787				
Db	3601	AGAGAGATGATCCCAAGGCCAGCAGGCACTTCCAGCTTACGCTTACGCTTAC	3657	QY	2525	GTACCCCTCTGACCCATTACCTGAACTTAGGAGAGACATCTGCTGCTA	2584	Db	4788	GTACCCCTCTGACCCATTACCTGAACTTAGGAGAGACATCTGCTGCTA	4847				
QY	1386	ACGAAATCGTACATGGAGATGGCTGAGACTTGGCTGACATTTGGCTGCTTASAA	1445	QY	2585	CCAAGGTTAGTAACCATCCATAACTCTGCTGAGCTAATTCGCTGAGGATGAGG	2644	Db	4848	CCAAGGCTCTGAACTCATGCTGAGCTTGGGACACAGAGGAATATGGAGG	4900				
Db	3658	ACGAAATCGTACATGGAGATGGCTGAGACTTGGCTGACATTTGGCTGCTTASAA	3716	QY	3776	-GGGAGAGGAAAGAAATCTCTACTGCTCTTCTGAGAGCTTGGGAGGCTTACGAGG	3834	QY	2645	CAAGAGACCGTGAAGGCAACAGGCTTAACTAACGATTCCACTCTTCTTACCAAGGCA	2704	Db	4901	TACAGAACTTACAGAGGAAAGGCTTAACTAACGATTCCACTCTTACCAAGGCA	4958
QY	1446	GGCTAAAGGAAGACTASAGAAARTCTTAYGATTACTCTGATGTCACCTAC	1505	Db	3835	CTGCTCTCTGACTGACTCTTCTGAGGCAACTAACTTAAAGGTAAGTAA	3894	QY	2705	CCTGCTCTTCTTATACCACGCTTACCTGAGGCTTATACCTGAGGTTAA	2764	Db	4959	TGTGTTCTCTATACCGCTGAGGCTTACCTGAGGTTAA	5018
Db	3717	GGACTAAGGAAACTCTAGGAAGAGCTTACATGATGTTACATGTCACCATATACA	3775	QY	1506	GGGGAGAGGAAAGAAATCTCTACTGCTCTTCTGAGAGCTTGGGAGGATGAGG	1565	QY	2765	GGAACGAGAGGTTTACASTCTGAGCTTGGCTTCTCTGAGGAGG	2824	Db	5019	GGAACGAGAGGTTTACASTCTGAGCTTGGCTTCTCTGAGGAGG	5078
QY	1566	GTGCTCTCTGACTGACTCTTCTGAGGCAACTAACTTAAAGGTAAGTAA	1625	Db	3895	CTCAGTAGCTGAGGAAAGGAAACGAGTAACTTAAAGGAAAGGAAAGGAAAGGAA	3954	QY	1745	AGCGGAAACAGACAACAGGGTTAAAGGAAAGGCCACCGCTTACCTGAGG	1804	Db	5079	TCCTGACTCTCAATTCTGTCAGGTTGCTTGTGAGAACTCTCAACCCARCATCTCACTCAC	2884
Db	3895	CTCAGTAGCTGAGGAAAGGAAACGAGTAACTTAAAGGAAAGGAAAGGAAAGGAA	3954	QY	1805	CAAGTGGACTTGGGAGCTCTGAGGAAAGGAAAGCTGAGGAAATGCTTAATAG	1864	QY	2885	CTGGACTRTTACCCAAAGGGTCAAGGAGTAGYCCCATTTTGCCAGGATAG	2944	Db	5139	CTGGACTRTTACCCAAAGGGTCAAGGAGTAGYCCCATTTTGCCAGGATAG	5198
Db	4015	AGGCGAACAGACAACAGGGTTAAAGGAAAGGCCACCGCTTACCTGAGG	4074	QY	1865	GCTGTGCTCTGCTGCTCAAGCAACTTAAAGGTTGTCAGAGAGTAG	1924	QY	2945	CCAGACTCTGAGCTTGGCTTCTCTGAGCTTGGCTTCTCTGAGGATG	3004	Db	5199	CCAGACTCTGAGCTTGGCTTCTCTGAGCTTGGCTTCTCTGAGGATG	5256
Db	4075	CAAGTGGACTTGGGAGCTCTGAGGAAAGGAAAGCTGAGGAAATGCTTAATAG	4134	QY	4135	GCTGTGCTCTGCTGCTCAAGCAACTTAAAGGTTGTCAGAGAGTAG	4194	QY	3005	ACTTTRGCVCCURTTCAAGAACCTTGGCTTCTCTGAGCTTGGCTTCTCTGAG	3064	Db	5257	ACTTTRGCVCCURTTCAAGAACCTTGGCTTCTCTGAGCTTGGCTTCTCTGAG	5316
QY	1925	CGCCCTCTCTGCTCTGCTCAAGCAACTTAAAGGTTGTCAGAGAGTAG	1984	Db	4195	CGGCCCTCTGCTCTGCTCAAGCAACTTAAAGGTTGTCAGAGAGTAG	4254	QY	3065	CCTCGYACCTGCTCACWGGTTCACASARAGTCACCTGCTCACAGG	3124	Db	5317	CCTCGYACCTGCTCACWGGTTCACASARAGTCACCTGCTCACAGG	5376
QY	1985	GGACAAAGGTGTTGAGGAGCAACTACCACTGAGTACAGGAGACTGAGG	2044	Db	4255	GGACAAAGGTGTTGAGGAGCAACTACCACTGAGTACAGGAGACTGAGG	4314	QY	3125	TAATACTTAGGCTAATATCCAAAGGCCACGARGGCCCTAGTGAGGAA	3184				

Db	5377	-----TACTTAGGGCTAAATTATCCAAAGGACCAAGGCCCTCAGTCAGGACACATCCA	5432	Db	6436	TTATGGCTTACAGAACCTAAGAGGAGTGGCAGCTTACITCGGGGTCTACAGAAA	6495
Qy	3185	GCCTATACTGGCTTATCCTCATCYCAAACCTTAAGCAATAAGGRRTCTGGCRT	3244	Qy	4265	GGAAGRAMAGGGAATASAGRGAAAYGCGAACAKATATGAGMARAAGAGCTGCA	4324
Db	5433	GCCTATACTGGCTTATCCTCATCCATCCAAACCTTAAGCAATAAGGRRTCTGGCRT	3244	Db	6496	GGAAGGAAAGGGAATAGAGAGACTGCAAGAGATATTGAGCACAAGAGCTGCA	6555
Qy	3245	RAYAGGYTCTCGCGAAWAT3EATTCCAGSTWTGCRRAATAGCCAGGYCATIWATA	3304	Qy	4325	AGGAGGACCTCATTAGAATGCTTAACTCCTTAGATAGGCTTAATCCCTTC	4384
Db	5493	RATAGGITCTCCGGAATATGGATTCCAGSSTATGGGAATAGCGAGGCTTAAATA	5551	Db	6556	AGGAGGACCTCATTAGAATGCTTAAACAGACCCCTAGTAGATAGGPATACTCCCTC	6615
Qy	3305	CASTAATTAAGGAAACTCAGAAAGCCAATACCCATTARTAGATSGAYAMCTGAAGYMR	3364	Qy	4385	CGGAAACCAAGGCCAGTACTCAGCAGAGAAACAGAATGGGAACCTCAGGAG-CAG	4443
Db	5612	AGTGGCTTCAGGCCCTTAAGAGGCTTAACCAAGGCCAGTGTTAGYTGCC	3424	Db	6616	CGGAAACCAAGGCCAGTACTCAGCAGAGAAACAGAATGGGAACCTCAGGAG-CAG	6675
Qy	3485	TTCACACGRTCCRAGGAGYGAESTGRACRYTGSRCYRACCTGASTAGGAAYTGT	3484	Qy	4444	TTTCCTCCCTGGAGCGGTAGCCACTGAGAATCTTGCTGCAACTAT	4503
Db	5717	TTCACACGATCCGAGGGATGAGCTGCAACTCTGGAATACCTGAGATGGAGATTGT	5716	Qy	4585	TTCATGCCAACTCATTTCTATAGTCACAG-AAGAACAGGAATAGCTTAGAGTCC	5716
Qy	3425	ACRGESCAAACACTTTSTTYATVTCAGAGAAAGACAGRAAGYAGCTTRGAGTCC	3484	Db	6676	TTTCCTCCCTGGAGGGCTAGCCACTGAGAAGGAAATCTTGCTGCAACTAT	6735
Db	5658	ACAGGCGAAAGACTTCTCATATGTCACAG-AAGAACAGGAATAGCTTAGAGTCC	5716	Qy	4594	TCARATGCCAACTCATTTACTGAGCAGGAAACCTTCAACTATCACGAGATATT	6735
Qy	3485	TTCACACGRTCCRAGGAGYGAESTGRACRYTGSRCYRACCTGASTAGGAAYTGT	3484	Db	6796	TCAGATGCCAACTCATTTACTGAGCAGGAAACCTTCAACTATCACGAGATATT	6855
Db	5717	TTCACACGATCCGAGGGATGAGCTGCAACTCTGGAATACCTGAGATGGAGATTGT	5716	Qy	4624	CAGGGCTGTCAGKTCGCCRARAATTAATCCTGCTTCCTVATCGCAAGCTCTGAG	6795
Qy	3545	TAGTGGCAAGGGTGRCYTATGTTAYGGTAGGGTGGCAGTACGAGTAGYKTGAT	3604	Db	6856	CAGGGCTGTCAGKTCGCCRARAATTAATCCTGCTTCCTVATCGCAAGCTCTGAG	6915
Db	5777	TAGTGGCAAGGGTTGACCTGTTACGAGTCTAGTGTGTCAGTAGCAGTGTAGT	5836	Qy	4684	ARACACARAACAGGCCATTACCGTGTGAGTACCTGCAACTCTGAGCTAACCC	4743
Qy	3605	CTGAGGCACTTAATAATAACGGAGAGATCTGTTACGAGTCTAGTGTGTCAGTAA	3664	Db	6916	AAACTCTGGGTTCTGTTACGAGTCTAGTGTGAGTACCTGTCAGGAGCTAACCC	4743
Db	5837	CTGAGCAGTTAAATAATAACGGAGAGATCTGTTACGAGTCTAGTGTGTCAGTAA	5896	Qy	4804	GCCTCCCTGTGAGGAGAAAGGCCAGAGGTTAAAGGCACTAGTCTGAGTAA	4863
Qy	3665	AATGGCTCTTACTGTAAGGGCACTGCTGCACTTAACTTACCCAG	3724	Db	7036	GCCTCCCTGTGAGGAGAAAGGCCAGAGGTTAAAGGCACTAGTCTGAGTAA	7095
Db	5897	AATGGCTCTTACTGCTAAGGAGCTGTGCTGTCAGCAACTGTTACTAATGTC	5956	Qy	4864	ATCCCCGATTCGACTTCCCGAGGCTTACAGAGTCAATAGCCCTGCTTCCAGGCC	7035
Qy	3785	YCNCATTCTCCAGACAACTGAAAGMTARAYATCTGCAACRATTAACTTC	3844	Db	7095	ATCCCCGATTCGACTTCCCGAGGCTTACAGAGTCAATAGCCCTGCTTCCAGGCC	7155
Db	6017	CCAGCTTCTCCAGACAACTGAAAGAATAAACATACGTCAACAGTAACTTC	6076	Qy	4924	ACGTAACCCAGGGAGATCCCGAGGCTTAGGTTAGGATAGATACACTACGCTGCA	4983
Qy	3725	AGCCTCTTACTGAAAGGCCAGTGTGCACTTGTCACCTTACCCAG	3784	Db	7156	ACGTAACCCAGGGAGATTCAGGCTTACGAGTATCTACGCTACGCTGCTGCA	7215
Db	5957	AAGCCTCTTACTGTAAGGGCACTGCTGCACTTAACTTACCCAG	6016	Qy	4984	AGGCCACGTCCTCAGGAGGAGGAAATGATGAAAYACTCAAGGACATCTAA	5043
Qy	3785	YCNCATTCTCCAGACAACTGAAAGMTARAYATCTGCAACRATTAACTTC	3844	Db	7216	AGGCCACGTCCTCAGGAGGAGGTTAGGATAGATGAAACATCTAAAGGACATCTAA	7275
Db	6077	AAACCTTATGCACTGAGGGACCTTGTAGGGTCTTGTACTGATCCGACC-TCAAC	6135	Qy	5044	AAGCAGCCAGGAAACCCACCTCAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCA	5103
Db	6017	CCAGCTTCTCCAGACAACTGAAAGAATAAACATACGTCAACAGTAACTTC	6076	Db	7276	AAGCAGCCAGGAAACCCACCTCAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCA	7335
Qy	3905	TGTTAATCTGTTGGAGTTCTGTGAGAAAGACTTGAAAGAAGYGGGTGAGCAG	3964	Qy	5104	ATCTGCCACTTCCCGAGGAGGACTTAGCCATCGAAATGCTGTPATGGAGGCC	5163
Db	6136	TTGTTAATCTGTTGGAGTTCTGTGAGAAAGACTTGAAAGGTTCTGAGCAG	6195	Db	7336	ATCTGCCACTTCCCGAGGAGGACTTAGCCATCGAAATGCTGTPATGGAGGCC	7395
Qy	3965	GTCACTGATATGGAAATYTGAAGAATATCCCTCACTCAGGACTAGTGTGAGT	4024	Qy	5164	TTCATACCATTCGCCATAGCTGAGGAGGAACTTACGCTGCTGCTGCTGCTGCTGCA	5223
Db	6196	GTCACTGATATGGAAATYTGAAGAATATCCCTCACTCAGGACTAGTGTGAGT	6255	Qy	5284	TTCATACCATTCGCCATAGCTGAGGAGGAACTTACGCTGCTGCTGCTGCTGCTGCA	5343
Qy	4025	GCAGAGCTTAAGCCYTCAYTKGGCACTAGTGTGAGGAGGAAAGGGYATATA	4084	Db	7516	TTCATACCATTCGCCATAGCTGAGGAGGAACTTACGCTGCTGCTGCTGCTGCTGCA	5343
Db	6256	GCAGAGCTTAAGCCYTCAYTKGGCACTAGTGTGAGGAGGAAAGGGYATATA	6315	Qy	5224	CCTTAGCCAAATTCACAGCTTAACTACGCTGCTGCTGCTGCTGCTGCTGCTGCA	5283
Qy	4085	TATCAGACTCTTACGTTACGTTACGTTACGTTACGTTACGTTACGTTACGTT	4144	Db	7456	CCTTAGCCAAATTCACAGCTTAACTACGCTGCTGCTGCTGCTGCTGCTGCTGCA	7515
Db	6316	TATCAGACTCTTACGTTACGTTACGTTACGTTACGTTACGTTACGTTACGTT	6375	Qy	5284	GAAGAACTATCCACCCCTGAGGAGGAACTTACGCTGCTGCTGCTGCTGCTGCA	5283
Qy	4145	AGGAATTCCTACTCYGGRGAACTCATCACATCGGAGGAACTTACGTTACGTT	4204	Db	7516	GAAGAACTATCCACCCCTGAGGAGGAACTTACGCTGCTGCTGCTGCTGCTGCA	7515
Db	6376	AGGAATTCCTACTCYGGRGAACTCATCACATCGGAGGAACTTACGTTACGTT	6435	Qy	7516	GAAGAACTATCCACCCCTGAGGAGGAACTTACGCTGCTGCTGCTGCTGCTGCA	7515
Qy	4205	TTAYTGGCWWTCACAGAACCTTACAGAGGTTGAGGTTGAGAGCTTACGTCYGGGCTCATCANA	4264	Db			

QY	5344	CATCCCTAGATACTCTGGGAAAGGACCCCTACCCAGTCATTATYTACCCCAACTGGG	5403	QY	6424	GGCTCTCGAGAATCTAGTGCTTCTCTCATTTCTAGTGCCCCYATGCCATCTACACT	6483
Db	7576	CATCCTAGATACTCTGGGAAAGGACCCCTACCCAGTCATTATCTACCCCAACTGGG	7635	Db	8655	GGCTCTCGAGAATCTAGTGCTTCTCTCATTTCTAGTGCCCCYATGCCATCTACACT	8714
QY	5404	TTAAGTGTGGTGGAGTGGACTACATCACCTGAGTCATTAATCTGGATACCTGC	5463	QY	6484	GACAATTTACAGATATGCTCATCTAAGCCCCAACAAAGTACCATCT	6543
Db	7636	TTAAGTGTGGTGGAGTGGACTACATCACCTGAGTCATTAATCTGGATACCTGC	7695	Db	8715	GRACAGATTATACTACATATGCTATCTAAGGCCAACAAAGTACCATCT	8774
QY	5464	CAAGGAACCTGAAATCGGGAGAACCTAGTATCTCTGTGAACTCTAGAGATT	5523	QY	6544	CTTTTGTATAGGAGGAGGTGCTGGTACTAGTAGCTACTGGCATTTGGATACCG	6603
Db	7696	CAAGGAACCTGAAATCGGGAGAACCTAGTATCTCTGTGAACTCTAGAGATT	7755	Db	8775	CCTTTGTATAGGAGGAGGTGCTGGTACTAGTAGCTACTGGCATTTGGATACCG	8834
QY	5524	TGCCCTGCTTCAACACAAACAGGAGAACCTAGTATCTCTGTGAACTCTAGAGATT	5583	QY	6604	ACCTCTACTCAGTCTACTACACTACTCTGAGTCATCTACCTGCA	6723
Db	7756	TGGCTCTCTCAACACAAACAGGAGAACCTAGTATCTCTGTGAACTCTAGAGATT	7814	Db	8895	GGCGACTCCCTGTCACCTGCAAGATCAACTTCTAGGAGAT	8954
QY	5584	GSCCTCCCTATCATATTCTCTKTGSTTACCTTCACTCTACCTGCA	5643	QY	6784	GAGAATGCTTAGTGTGCTTACCTTCACTCTACCTGCA	6843
Db	7815	GCCCTCCCTATCATATTCTCTKTGSTTACCTTCACTCTACCTGCA	7874	Db	9115	GAGAATGCTTAGTGTGCTTACCTTCACTCTACCTGCA	9074
QY	5644	CCCCCTCATGGCTGTATGACAGTGGCTCCYACCMAGATTCTATGAGATG	5703	QY	6724	ATCGAGAGCTTAGTGTGCTAAGCTGARAGGGGAACTCTTATTTAGG	6783
Db	7875	CCCCCTCATGGCTGTATGACAGTGGCTCCYACCMAGATTCTATGAGATG	7934	Db	9115	AATCGAGAGCTTAGTGTGCTAAGCTGARAGGGGAACTCTTATTTAGG	9074
QY	5764	TTCAGTCGCCAACCCATATGCCCAATGCTATAGGAGTCTTSTAAGGGAAACCCACC	5763	QY	6844	CSAGATGTAACAGTAKAGCAGCAGACTCTAGGAGTCTGCTGGAGACTGGA	6903
Db	7995	TTCAGTCGCCAACCCATATGCCCAATGCTATAGGAGTCTTSTAAGGGAAACCCACC	8054	Db	9075	CGAGATGTAACAGTAKAGCAGCAGACTCTAGGAGTCTGCTGGAGACTGGA	9134
QY	5824	GERAACTACTCATATTGACAGAAMATGATTAACTCTAGTGTCTGGAGACTGGA	5883	QY	6904	ACCRATGGATCCCTGATCTCCCTCTAGGACTCTGCTTAACCGCTGAAAGGGGAACTCTTATTTAGG	6963
Db	8004	GERAACTACTCATATTGACAGAAMATGATTAACTCTAGTGTCTGGAGACTGGA	8114	Db	9135	ACCCAGTGGATCTCCCTCTAGGACTCTGCTTAACCGCTGAAAGGGGAACTCTTATTTAGG	9194
QY	5884	GTCACTGTGTGGACTACTCTACCCAACTGGTAGTCTATGGGGTGGAGTC	5943	QY	6964	CTCCCTCTGGACCTGCTTCTCTCTGTTACTTGTCCTTCAGAATGAA	7023
Db	8115	GTCACTGTGTGGACTACTCTACCCAACTGGTAGTCTATGGGGTGGAGTC	8174	Db	9195	CTCCCTCTGGACCTGCTTACTTGTCCTCTGTTACTTGTCCTTCAGAATGAA	9254
QY	5944	GATCAGGGAAAGAAACATGTAAGAAGTAACTCCAACTCACCGGAACTAGGC	6003	QY	7024	GCTGTRAACTACAAATGGGCCAGATGCTGAACTAGTAACTGCTACCGAGCCC	7083
Db	8175	GATCAGGGAAAGAAACATGTAAGAAGTAACTCCAACTCACCGGAACTAGGC	8234	Db	9255	GCTGTRAACTACAAATGGGCCAGATGCTGAACTAGTAACTGCTACCGAGCCC	9314
QY	6004	ACCTCTAGCCCTACAAAGGGACTAGATCTCAAACCTCTGTTACCAT	6063	QY	7084	CTGGACGGCTGTGAGCCAGATGCTGTTAACTGCTCAAGGCCACCCCTCTGAG	7143
Db	8235	ACCTCTAGCCCTACAAAGGGACTAGATCTCAAACCTCTGTTACCAT	8294	Db	9315	CTGGACGGCTGTGAGCCAGATGCTGTTAACTGCTCAAGGCCACCCCTCTGAG	9374
QY	6064	ACTCGCTCTGGTAACTTATAACCCCTACTGGCTCATGGGTCTGGCCAA	6123	QY	7144	GAATTCAGTCAGAACCTCTACTACGGCCAACTCAGGAGACAGTAGGGG	7203
Db	8295	ACTCGCTCTGGTAACTTATAACCCCTACTGGCTCATGGGTCTGGCCAA	8354	Db	9375	GAATTCAGTCAGAACCTCTACTACGGCCAACTCAGGAGACAGTAGGGG	9434
QY	6124	ACCCCTACTACTGTGGATATGCCCTGAACTTCAGGCTATGTTCACTCC	6183	QY	7204	SCTGCCCAACCTCCCAACAGCAGACTCTAGGTTCTCTGTTAGGAGATGGGGACTGAGAC	7263
Db	8355	ACCCCTACTACTGTGGATATGCCCTGAACTTCAGGCTATGTTCACTCC	8414	Db	9435	CATGCCCAACCTCCCAACAGCAGACTCTAGGTTCTCTGTTAGGAGATGGGGACTGAGAC	9494
QY	6184	GTACCTGACACATGGGAACTTCAGCACAAATAACACACTTCGGTTAGGAA	6243	QY	7264	AGGACTAGCTGGATTCTCTAGGCTGATAGAATCCYTAAGCTGTTAGGAGATGGGGACTGAGAC	7323
Db	8415	GTACCTGACACATGGGAACTTCAGCACAAATAACACACTTCGGTTAGGAA	8474	Db	9495	AGGACTAGCTGGATTCTCTAGGCTGATAGAATCCYTAAGCTGTTAGGAGATGGGGACTGAGAC	9554
QY	6244	CCTCTGTTCAATSTGGAAATAACCCATACCTCAACCTACCTGTMATTAGC	6303	QY	7324	ACATCCACCTTAACACGGGGCTGCACTTAGTCACCTGACCAATCAGAGCTC	7383
Db	8475	CCTCTGTTCAATSTGGAAATAACCCATACCTCAACCTACCTGTMATTAGC	8534	Db	9555	ACATCCACCTTAACACGGGGCTGCACTTAGTCACCTGACCAATCAGAGCTC	9614
QY	6304	ATACTACATACACACCACTCCAGGAACTTCAGGAGTGGACTCTCCACACAA	6363	QY	7384	ACTAAATGTTAACTAGGCAAGACAGGAGGTTAGGAAATAGCCATCATTTATGCGMTG	7443
Db	8535	ATACTACATACACACCACTCCAGGAACTTCAGGAGTGGACTCTCCACACAA	8594	Db	9615	ACTAAATGTTAACTAGGCAAGACAGGAGGTTAGGAAATAGCCATCATTTATGCGMTG	9674
QY	6364	GTCTGCCTACCTCAGGAACTTCAGGAGTGGACTCTCCACACAA	6423	QY	7444	AGAGCAGAGGAGGAGTGGAGTAACTCCAGGAGTGGACTCTCCACACAA	7503
Db	8595	GTCTGCCTACCTCAGGAACTTCAGGAGTGGACTCTCCACACAA	8654	Db	9675	AGAGCAGAGGAGGAGTGGAGTAACTCCACACAA	9734
QY	7504	CAACCCCTTGGCTCCCTCCCTTGTATGGAGCTCTTTCTAGCTTACTCT	7563				

Db	3432	TCTCCAGACTTAAGCAATAAACAGACTTAGTAAATTCTCAGATAACCTGTATGG	3491	Qy	2225	GTCACTGATACTTYGCCAGCCACTAAGTATGACTGEGGAGCTTATTCCTTCAC	2284
Qy	1146	CATATGTTGTTTACAGGGTTAGGAACATCTTGATCTGACATGGAGATATA	1205	Db	4559	GTCACTGATACTTCCTCCAGCCACTAAGTATGACTGEGGAGCTTATTCCTTCAC	4616
Db	3492	CTTATGTTGTTTACAGGGTTAGGAACATCTTGATCTGACATGGAGATATA	1205	Qy	2285	ATCCTTCTTAATTATCCTGAAAGCCCACTACCTTCTTGGAGAGACATCTTCA	2344
Qy	1206	TGCACTGCTTAATCAGACACTAACCCRAATGAGAGAGTGCACCATACTTGATGAGATA-A	3550	Db	4617	ATGCTTCTTAATTATCCTGAAAGCCCACTACCTTCTTGGAGAGACATCTTCA	4676
Db	3551	TGCACTGCTTAATCAGACACTAACCCRAATGAGAGAGTGCACCATACTTGATGAGATA	3610	Qy	2345	AAGCAGGGCCATTATACTGAGACTAGGAGATANGGATGACAACAGAGG	2404
Qy	1266	GAGTTTGGGATCTCGTTACTCAGTCAAGGCAAGGATGATGGAGATGGAGG	1325	Db	4677	AAGCAGGGCCATTATACTGAGACTAGGAGATANGGATGACAACAGAGG	4735
Db	3611	GAGAGTTGGGATCTCGTTACTCAGTCAAGGAGATGGAGATGGAGG	1367	Qy	2405	CTTGAGGAGAGGATTAATCTCTGAACTGAGTGGGACAGCCATCTGGGAC	2464
Qy	1326	AAAGAAATGATTCCTCACAGGCGACGARGAGTCCAGCTTCTGAGACCTATGG	1385	Db	4736	CTTGAGGAGAGGATTAATCTCTGAACTGAGTGGGACAGGAGACATGGGAG-CAG	4794
Db	3668	AAAGAGATGATTCCTCACAGGCGACGARGAGTCCAGCTTCTGAGACCTATGG	1385	Qy	2465	GAATGCCGCTCTGTCAGTAACTTAAGTGGGACAGAGCAATTGGACAGGCCAA	2524
Qy	1386	ACAGAAATGATTCCTCACAGGCGACGARGAGTCCAGCTTCTGAGACCTATGG	1445	Db	4795	GAATGCCGCTCTGTCAGTAACTTAAGTGGGACAGAGCAATTGGACAGGCCAA	4854
Db	3725	ACACAGAAATCAGACAT-GGAGATGGCTGGAGACATTTGCACTTGCTGAGCTGAGA	3783	Qy	2475	GTACCCCTCAGACCAAGGCCACAGGATCCAAAGATGTTAGGACTAAAGC	2584
Qy	1446	GGACTAAGAAACTASGAAARTCTGAAATTACTAGATGTCACCTAACAA	1505	Db	4885	GTACCCCTCAGACCAAGGCCACAGGATCCAAAGATGTTAGGACTAAAGC	4914
Db	3784	GGACTAAGAAACTASGAAARTCTGAAATTACTAGATGTCACCTAACAA	1505	Qy	2555	CCAGGGCTTAGAAACATGCTTAACCTGAGTGGGAGACTTCCACTCTTCCACAAAGCA	5025
Qy	1506	GGGGAAAGGAAAGAAATCTACTGCTTCCTTCTGGAGAGACTAACAGGAGG	1565	Db	4915	CCAGGGCTTAGAAACATGCTTAACCTGAGTAAACCCCTCAGACTCC-----AATTAGGAG	4967
Db	3843	-GGAGGGAGAGAAATCTACTGCTTCCTTCTGGAGAGACTAACAGGAGG	1591	Qy	2645	CACAGAACCCCTGAGCAGTGGAGGAGCTCAAAGATGTTAGGACTAAAGC	2704
Qy	1566	GTCGCTCTGTCACCTGACTCTCTGAGGCACTAACTTAAAGCTTAAGCTTAAGCTTAAGCTTAC	1625	Db	4968	TACGAAACCCACAGACAGTGA-GTTAGTCAACATCTCAGGATAT-GAGC	5025
Db	3902	GTCGCTCTGTCACCTGACTCTCTGAGGCACTAACTTAAAGCTTAAGCTTAC	1625	Qy	2705	CGTGTCTTACCCAGCTGAGTAAACATGCTTAACCTGAGTAAACCCCTCAGACTCC-----AATTAGGAG	2764
Qy	1626	CTGACTGAGCTGAGACATTGAAAAGCTCAAGTGGTAGGAGGAACTGAGC	1684	Db	5026	TGTGTCTCTAGACGACTGAGTACTCTAGGCTTATACTCTGTTCCAAATACAGA	5085
Db	3962	CTGACTGAGCTGAGACATTGAAAAGCTCAAGTGGTAGGAGGAACTGAGC	1684	Qy	2765	GGAGCAGCTGAGTGGTTAGCTGGACCTTCTGGATTCAGGAACTCTCAGGATAT-GAGC	2824
Qy	1685	ACTTAAACCCATTGAACTGTTGAGCTTAAATAGAGCTTACGGAGGAC	1744	Db	5086	GGAGCAGCTGAGTGGTTAGCTGGACCTCTGGAGGACCTTCATCCATCGTAC	5145
Db	4022	ACTTAAACCCATTGAACTGTTGAGCTTAAATAGAGCTTACGGAGGAC	1744	Qy	2825	TCCGTACGCTCAATTCTGTTGCTGAGTACTCTCAACCACATCTCACTCAC	2884
Qy	1745	AGGGAAACAGGAAACAGGAACTTAAAGGACCCCTTAGTCAGACCTCAGG	1804	Db	5146	TCCGTACGCTCAATTCTGTTGCTGAGTACTCTCAACCACATCTCACTCAC	5205
Db	4082	AGGGAAACAGGAAACAGGAACTTAAAGGACCCCTTAGTCAGACCTCAGG	1804	Qy	2885	CTGGACTTTACCCAGGGTCAGGATGTCAGGATGTCAGGATGTCAGGATG	2944
Qy	1805	CAAGTGGACTTTGAGGCTCTGGAAAGGAAAGCTGGCAATTGAGTGCCTAATGG	1864	Db	5205	CTGGACTATTACCCAGGGTCAGGATGTCAGGATGTCAGGATGTCAGGATG	5265
Db	4142	CAAGTGGACTTTGAGGCTCTGGAAAGGAAAGCTGGCAATTGAGTGCCTAATGG	1864	Qy	2945	CCAGACTCTGAGCCTTACCTGAGGAACTCTGGAGATGTCAGGATGTCAGGATG	5326
Qy	1865	GCTGCTTCCAGTGGGTTAACGAGACTTAACGAGACCTTAAAGGAGCTGGCAATTGAGTGCCTAATGG	1924	Db	5266	CCAGACTCTGAGCCTTACCTGAGGAACTCTGGAGATGTCAGGATGTCAGGATG	5323
Db	4202	GCTGCTTCCAGTGGGTTAACGAGACCTTAAAGGAGCTGGCAATTGAGTGCCTAATGG	1924	Qy	3005	ACTTTCAGCTGCTGAGGAACTCTGGAGATGTCAGGATGTCAGGATGTCAGGATG	3064
Qy	1925	CCGCCTCTGTCATGCCCCTATTTCAGGAACTCTGGAGGAACTCTGGAGG	1984	Db	5324	ACTTTCAGCTGCTGAGGAACTCTGGAGGAACTCTGGAGGAACTCTGGAGG	5393
Db	4262	CCGCCTCTGTCATGCCCCTATTTCAGGAACTCTGGAGGAACTCTGGAGG	1984	Qy	3065	CCTCGCYACTCTGGCTCACWGGTTCAAACSRARGCTCACTCTGTCAGGAGG	3124
Qy	1985	GGACAAAGCTTCTGGAGGAACTAACCCACTAACCCACTAACCCACTAACCC	2044	Db	5384	CCTCGCYACTCTGGCTCACWGGTTCAAACSRARGCTCACTCTGTCAGGAGG	5443
Db	4322	GGACAAAGCTTCTGGAGGAACTAACCCACTAACCCACTAACCCACTAACCC	2044	Qy	3125	TAATCTTAGRCRARATATCCAAAGGCCACRGCCCTCTGGAGGAACTCTGG	3184
Qy	2045	TGCTGGGAGGCCATCCATGCCATCCATGCCATCCATGCCATCCATGCCATCC	2104	Db	5444	--TCTAGGGCTTAATTATCCAAAGGCCACGGCCCTCTGGAGGAACTCTGG	5499
Db	4382	TGCCCTGGGAGGCCATCCATGCCATCCATGCCATCCATGCCATCCATGCCATCC	2104	Qy	3185	GCCTATCTGGCTTATCCCTCATCCAAACCCCTAACGAACTAAGRGRITCTGCT	3244
Qy	2105	ATGAGGGAGGAGGTTCTCTGGAAACTGGCGCTCTTAGCTTACTCTC	2164	Db	5500	GCCTATCTGGCTTATCCCTCATCCAAACCCCTAACGAACTAAGRGRITCTGCT	5559
Db	4442	ATGAGGGAGGAGGTTCTCTGGAAACTGGCGCTCTTAGCTTACTCTC	2164	Qy	3245	AAYGGYTCTGGAGAATAGCCAGGWTGGERAAATAGCCAGYCA-TAWATA	3304
Qy	2165	TGTCCGGAGAACCTCTCAGATCTGTTACTATCTGGGGGGTCNTAGACGGCA	2224	Db	5560	AATAGGTTCTGGCGAAATGGAT-CCCAAGGTTAGGCGAAATAGCCAGTCATTAATA	5618
Db	4501	TGTCCGGAGAACCTCTCAGATCTGTTACTATCTGGGGGGTCNTAGACGGCA	2224				

QY	3305	CASTAATTAGGAAACTCAGAAAGGCCATACCCATTARTAAGGATGGAYAMTCGAGYMR	3364	Db
Db	5619	CACTAATTAGGAAACTCAGAAAGGCCATACCCATTARTAAGGATGGACACTGGAGTAG	5678	Db
QY	3365	AAGTGCTTCCAGGCCCTAACGAAAGGCCATACCCATTARTAAGGATGGCTTCAGTGA	3424	Db
Db	5679	AGTGCCTTCCAGGCCCTAACGAAAGGCCATACCCATTARTAAGGATGGCTTCAGTGA	5724	Db
QY	3425	AACTGGCAAGACTTTSVATYTCAGAAAGGCCATACCCATTARTAAGGATGGACACTGGAGTAG	3484	Qy
Db	5725	AACTGGCAAGACTTTSVATYTCAGAAAGGCCATACCCATTARTAAGGATGGCTTCAGTGA	5783	Db
QY	3485	TAAACAGTCRAGGAYGAGCTTCACCCAGTAAAGGAAATCTGAGTAAAGGAAATCTGAGT	3544	Qy
Db	5784	TAAACAGTCRAGGAYGAGCTTCACCCAGTAAAGGAAATCTGAGTAAAGGAAATCTGAGT	5843	Db
QY	3545	TAGTGCAAAAGGGTGRCYCATGTTAYGGGTAGGGGCACTAGCTGAGTYKPTAGT	3604	Qy
Db	5844	TAGTGCAAAAGGGTGRCYCATGTTAYGGGTAGGGGCACTAGCTGAGTYKPTAGT	5843	Db
QY	3605	CTGAGCAAGGAACTTCTTAACTACAGGAGATCTACTGTGACATCTCATGAKGTGA	3664	Qy
Db	5904	CTGAGCAAGGAACTTCTTAACTACAGGAGATCTACTGTGACATCTCATGAKGTGA	5953	Db
QY	3665	ATYGCATACTCACTCTAACAGGAGACTTGCGCTCTAGCAAACTGTTACTTAATRTC	3724	Qy
Db	5964	ATYGCATACTCACTCTAACAGGAGACTTGCGCTCTAGCAAACTGTTACTTAATRTC	6023	Db
QY	3725	AGGCTCTTAACTCTGAGRGCCAGTGTGCACTGTGCAACTGTTACTTAACCCAG	3784	Qy
Db	6024	AGGCTCTTAACTCTGAGRGCCAGTGTGCACTGTGCAACTGTTACTTAACCCAG	6083	Db
QY	3785	YCNCATTCTTCCGACAATGAGAAAGATARAAYATACTGCACARTAATTCTC	3844	Qy
Db	6084	YCNCATTCTTCCGACAATGAGAAAGATARAAYATACTGCACARTAATTCTC	6143	Db
QY	3845	AAACTATCCATTGAGGGACCTTYAGARGTCCTYTGACTGATCCYGACCTTCAC	3904	Qy
Db	6144	AAACTATCCATTGAGGGACCTTYAGARGTCCTYTGACTGATCCYGACCTTCAC	6202	Db
QY	3905	TRGTATACGTGGAGTCTCTGTAGAAAAGGACTTGAGAAGGGTAGCAG	3964	Qy
Db	6203	TRGTATACGTGGAGTCTCTGTAGAAAAGGACTTGAGAAGGGTAGCAG	6262	Db
QY	3965	GTCAGTATAATGGATAYTGAGTATCCCTCACCGAACACTGTGTYAGCTR	4024	Qy
Db	6263	GTCAGTATAATGGATAYTGAGTATCCCTCACCGAACACTGTGTYAGCTR	6322	Db
QY	4025	GGAGAAATATACCCYTCATKGCCACTAGAAATTAGGAAGAAGAAANGGAAATATA	4084	Qy
Db	6323	GGAGAAATATACCCYTCATKGCCACTAGAAATTAGGAAGAAGAAANGGAAATATA	6382	Db
QY	4085	TATAAGACTCTRARTATGCTYACTACTCTNTCCATGCCATGMRGAATATGSAAGAA	4144	Qy
Db	6383	TATAAGACTCTRARTATGCTYACTACTCTNTCCATGCCATGMRGAATATGSAAGAA	6442	Db
QY	4145	AGGGAACTCTRACTCTYAGRGACACCTATCAMACATCAGGAGCATTAGGARATA	4204	Qy
Db	6443	AGGGAACTCTRACTCTYAGRGACACCTATCAMACATCAGGAGCATTAGGARATA	6502	Db
QY	4205	TATYGGCGWGTACAGAAACTCTAGGAGGTGMMAGCTTACACTGCGYGGGTATCANAA	4264	Qy
Db	6503	TATYGGCGWGTACAGAAACTCTAGGAGGTGMMAGCTTACACTGCGYGGGTATCANAA	6562	Db
QY	4265	GGAAGRAAGGAAATASANGRGAAVTGCCAGCAKATATTGAGAGAAAGGCGCA	4324	Qy
Db	6563	GGAAGRAAGGAAATASANGRGAAVTGCCAGCAKATATTGAGAGAAAGGCGCA	6622	Db
QY	4325	AGCGAGACCTCCATTAGAATGCTTAAACTCTCTTAGTATAGGTTATCCTC	4384	Qy
Db	6623	AGCGAGACCTCCATTAGAATGCTTAAACTACCAACCCCTAGGTTATCCTC	6682	Db
QY	4385	CGGAACACAGGCCAGTACTCAGCGGGAGAACAGAATGGGAACCTCAGGG-CAG	4443	Qy
QY	6683	CGGAAACCAAGGCCAGTACTCAGCGGGAGAACAGAATGGGAACCTCAGGG-CAG	6742	Db
QY	4444	TTTCTCCCTCGGAGGGTAGGACTGAGAGAGGAAATACTTTGCTGCACTAT	4503	Qy
Db	6743	TTTCTCCCTCGGAGGGTAGGACTGAGAGGAAATACTTTGCTGCACTAT	6802	Db
QY	4504	CCATGGAATTACTAAACCTTCTACCAAACTTCATCTAGGATGACATGACACCA	4563	Qy
Db	6803	CCATGGAATTACTAAACCTTCTACCAAACTTCATCTAGGATGACATGACACCA	6862	Db
QY	4564	TCARATGCCIAATCATTTACTGACGACAGGGCTTCAAACATACGACARATAK	4623	Qy
Db	6863	TCARATGCCIAATCATTTACTGACGACAGGGCTTCAAACATACGACARATAK	6922	Db
QY	4624	CAGGCTGTGAAKGUTGCCRARAATAATCCCTGCTTATGCCAGCCTCTCAG	4683	Qy
Db	6923	CAGGCTGTGAAKGUTGCCRARAATAATCCCTGCTTATGCCAGCCTCTCAG	6982	Db
QY	4684	AAACAAARACAGGCATTACCTGCRARARACTGGCAACTGATTACCCACAAAGCC	4743	Qy
Db	6983	AAACAAARACAGGCATTACCTGCRARARACTGGCAACTGATTACCCACAAAGCC	7042	Db
QY	4744	AAACCTCAGGGATTCTAGTACTACTAGTGTGGTARATACTTCAGGGTTGGGCRAG	4803	Qy
Db	7043	AAACCTCAGGGATTCTAGTACTACTAGTGTGGTARATACTTCAGGGTTGGGCRAG	7102	Db
QY	4804	GCCTTCCCCTGTAGACAGAAAGGCCAGAGGATAATAGGCAACTAGTCTCATGAATA	4863	Qy
Db	7103	GCCTTCCCCTGTAGACAGAAAGGCCAGAGGATAATAGGCAACTAGTCTCATGAATA	7162	Db
QY	4864	ATTCAGAGATCGGACTTCCCGAGGCTTCAGGTGACATAGCCCTGTTCCAGGC	4923	Qy
Db	7163	ATTCAGAGATCGGACTTCCCGAGGCTTCAGGTGACATAGCCCTGTTCCAGGC	7222	Db
QY	4924	ACAGTAACGGGAGTATCCAGGGTACAGGCTTACAGAGTACAGTGTGAGTACCTG	4983	Qy
Db	7223	ACAGTAACGGGAGTATCCAGGGTACAGGCTTACAGAGTACAGTGTGAGTACCTG	7282	Db
QY	4984	AGSCCACAGCTCTAGGGAGGTGAGAATGATGAAATAGGCAACTAGTCTCATGAATA	5043	Qy
Db	7283	AGGCCACAGCTCTAGGGAGGTGAGAATGATGAAATAGGCAACTAGTCTCATGAATA	7342	Db
QY	5044	AAGCAAACCCAGGAAACCACCTCACATGGCTGTTGTGCTTATGGCTTAAGAA	5103	Qy
Db	7343	AAGCAAACCCAGGAAACCACCTCACATGGCTGTTGTGCTTATGGCTTAAGAA	7402	Db
QY	5104	ATCTGCAACTTCCCAAAGCAGGAGCTTACGCCATAGGAATGCTGTTGGAGGCC	5163	Qy
Db	7403	ATCTGCAACTTCCCAAAGCAGGAGCTTACGCCATAGGAATGCTGTTGGAGGCC	7462	Db
QY	5164	TCTAACTCATGCTGAGCTGACGACAGCTTACGCCATAGGAATGCTGTTGGAGGCC	5223	Qy
Db	7463	TCTAACTCATGCTGAGCTGACGACAGCTTACGCCATAGGAATGCTGTTGGAGGCC	7522	Db
QY	5224	CCTTAGGCAATTCACAGGAACTTACGCCATAGGAATGCTGTTGGAGGCC	5283	Qy
Db	7523	CCTTAGGCAATTCACAGGAACTTACGCCATAGGAATGCTGTTGGAGGCC	7582	Db
QY	5284	GAAAGAACTATCCACCCWWTGAGCATGTTAGTCAATGCTGCTTCCTCTAATTC	5343	Qy
Db	7583	GAAAGAACTATCCACCCWWTGAGCATGTTAGTCAATGCTGCTTCCTCTAATTC	7642	Db
QY	5344	CATCCCTAGATACCTCTGGAGGACCTACCTGAGCATTTATYACCCAACTGCGG	5403	Qy
Db	7643	CATCCCTAGATACCTCTGGAGGACCTACCTGAGCATTTATYACCCAACTGCGG	7702	Db
QY	5404	TAAAGTGCTGAGCTGAGCTGAGCTACACTGAGCTGAGCTGAGCTGAGCTGAG	5463	Qy
Db	7703	TAAAGTGCTGAGCTGAGCTACACTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	7762	Db
QY	5464	CGAAAGACCTGAAATTCAGGAGAACAGAATGGGAACCTCAGGG-CAG	5523	Qy

QY	5524	TGCCTGCTGCTCACACACACACAGGAGGAACTAACTATAATCCCCATG	5583	Db
Db	7823	TGCCCTGCTCTCGAACACACACAGGAGGAAGTAACTAAAT-CCCATG	7881	QY
QY	5584	GSCCTCCTTATCAATTTCCTCTTASTGTTSSTTYACCTTTCCTACTGCA	5643	Db
Db	7882	GCCCTCCCTATCATATTTCCTTACTGTTTACCTCTTACTGCA	7941	QY
QY	5644	CCCCTCCATGCCSTGTATGACCGTAGTCCCTYACMAGATTATGAGATG	5703	Db
Db	8062	TTCACTGCCAACACCATATGCCGCACTGCTACTCTGCACTTGTGATGAT	8001	QY
QY	5704	CAGGGTCCCGAATTATGATGCCATCGTATAGGAGCTTSTAGGACCCACC	5763	Db
Db	8002	CAGGTTCCCGAATTATGATGCCATCGTATAGGAGCTTSTAGGACCCACC	8061	QY
QY	5764	TTCACTGCCACACCATATGCCGCACTGCTACTCTGCACTTGTGATGAT	5823	Db
Db	8062	TTCACTGCCAACACCATATGCCGCACTGCTACTCTGCACTTGTGATGAT	8121	QY
QY	5824	GCAAATACATTATGGACAGGAAATGATTATCTCTAGTGGCCCTGAGGACTTGCA	5883	Db
Db	8122	GCAAATACATTATGGACAGGAAATGATTATCTCTAGTGGCCCTGAGGACTTGCA	8181	QY
QY	5884	GTCAGCTGCTGTGACTACTCCCAACTGTTAGCTGATGGGGTGGAGTCATG	5943	Db
Db	8182	GTCAGCTGCTGTGACTACTCCCAACTGTTAGCTGATGGGGTGGAGTCATG	8241	QY
QY	5944	GATCGGGCAAGAAACACTGTAAGAGTAATCTCCAACTACCGGGTACATGGC	6003	Db
Db	8242	GATCGGGCAAGAAACACTGTAAGAGTAATCTCCAACTACCGGGTACATGGC	8301	QY
QY	6004	ACCTCTAGCCCTACAAAGCTACATCTCTCAAAACTATGAAACCCCTCGPACCCAT	6063	Db
Db	8302	ACCTCTAGCCCTACAAAGCTACATCTCTCAAAACTATGAAACCCCTCGPACCCAT	8361	QY
QY	6064	ACTCCCTGGTAGCCATTATACACCCATGGGCTCATCGAGCTCGCCCCAA	6123	Db
Db	8362	ACTCCCTGGTAGCCATTATACACCCATGGGCTCATCGAGCTCGCCCCAA	8421	QY
QY	6124	AACCTACTACTGTGGATATGCCCTCCCTGAACCTCAGCCATATGTTCAATCC	6183	Db
Db	8422	AACCTACTACTGTGGATATGCCCTCCCTGAACCTCAGCCATATGTTCAATCC	8481	QY
QY	6184	GTACTGAACTGACACTTCGCCACAGAAACACCACTTCGTTTACTAGGA	6243	Db
Db	8482	GTACTGAACTGACACTTCGCCACAGAAACACCACTTCGTTTACTAGGA	8541	QY
QY	6244	CCTCTGTTCAASITGGAAATACCACTCAACCTCACTGTGAAATTAGC	6303	Db
Db	8542	CCTCTGTTCAATGGAAATACCACTCAACCTCACTGTGAAATTAGC	8601	QY
QY	6304	AATATCATACACACCAACTCCATGCACTGGGGTAACTCTCCACACAAATA	6363	Db
Db	8602	AATATCATACACACCAACTCCATGCACTGGGGTAACTCTCCACACAAATA	8661	QY
QY	6364	GTCCTGCTACCTCAGGAATTTCCTGCTGACTCTGCACTGTTGTTGAAT	6423	Db
Db	8662	GTCCTGCTACCTCAGGAATTTCCTGCTGACTCTGCACTGTTGTTGAAT	8721	QY
QY	6424	GGCTCTCAGATCTATGTCCTCTCATCTAGGCCCTATGRCCATCCTAC	6483	Db
Db	8722	GGCTCTCAGATCTATGTCCTCTCATCTAGGCCCTATGRCCATCCTAC	8781	QY
QY	6484	GRACAGAGATTATCACTATCTAGTATCTAGGCCAACAAAGAGTACCCATT	6543	Db
Db	8782	GRACAGAGATTATCACTATCTAGGCCAACAAAGAGTACCCATT	8841	QY
QY	6544	CCTTGTATGGAGGAGGACTCTGGTCACTGCTGCTGCTGCTGCTGCTG	6603	Db
Db	8842	CCTTGTATGGAGGAGGACTCTGGTCACTGCTGCTGCTGCTGCTGCTG	8901	QY
RESULT 14	AY101586	AV101586	10229 bp	DNA
Locus			linear	FRI 11-FEB-2004

DEFINITION	Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERWWEI
LOCUS	allele A, complete sequence.
ACCESSION	AY101586
VERSION	AY101586.1
KEYWORDS	
ORGANISM	Pan troglodytes (chimpanzee)
Bukarjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.	
REFERENCE	1 (bases 1 to 10229)
AUTHORS	Mallett,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnard,B., Lucoff,G., Duret,L. and Mandrand,B.
TITLE	The endogenous retroviral locus ERWWEI is a bona fide gene involved in hominid placental physiology
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBLISHED	14/7/2002
REFERENCE	2 (bases 1 to 10229)
AUTHORS	Mallett,F., Bouton,O. and Oriol,G.
TITLE	Direct Submission
JOURNAL	Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allee d'italie, Lyon 69364 cedex 07, France
FEATURES	
Source	
1..10229	
/organism="Pan troglodytes"	
/mol_type="Genomic DNA"	
/isolate="1"	
/isolation_source="PRMC oE chimpanzee 1"	
/db_xref="Taxon:9598"	
/cell_type="PRMC"	
/note="isolated by PCR; ERWWEI locus, allele A	
/endogenus_virus: ERV-W"	
1..779	
/join="248..810 2889..9820)	
/note="putative mRNA transcript 1; gag-pol pseudogene"	
/join="248..810 7592..9820)	
/product="envelope glycoprotein"	
/note="putative mRNA transcript 2"	
/join="248..810 9255..9820)	
/note="putative mRNA transcript 3; no coding region"	
7819..9435	
/note="syncytin"	
/product="envelope glycoprotein"	
/protein_id="PAM68165.1"	
/db_xref="GI:7544408"	
/translation="MALPHIFLFTVLLPSFTLAPPCCRMTSSPYQEFLNWPMQRPTVCMYFTHGMSGGPPDFTAHTRHPRNCASATLGHANTHYWCKTINPSRCRGIVLHTRLVSLFNTLIGIHEVSAGAATTCGAGCTTGTGAGATGTTGCGACATGTTTCAVLRGVPSNBLITHTNLTCYKPSNTDTNSOCRDTWPTQIOTCPLSPGAGFVGTSAVRCANGSSSSEMCPSRNPVSKWPRKPVLPFLPVGAGFVGA	
GNIAKPSYRPSKGGPPDFTAHTRHPRNCASATLGHANTHYWCKTINPSRCRGIVLHTRLVSLFNTLIGIHEVSAGAATTCGAGCTTGTGAGATGTTGCGACATGTTTCAVLRGVPSNBLITHTNLTCYKPSNTDTNSOCRDTWPTQIOTCPLSPGAGFVGTSAVRCANGSSSSEMCPSRNPVSKWPRKPVLPFLPVGAGFVGA	
TGCGAGTGTSTPQIYLSQBLNGDMERVAASLVLQDQLNLSAAVLONERALDILTAERGGTGFISKECCYVNQSGIVTVEKVERDRIRRRAEELRNTSPWGLISQMMWILPFLGLIAATILLGPGCIPNLYVNSRVEAKVLUQMPHQSKTKIVYRPLRPRPQRSDNDIKPPEPEELTAQPLRPNSS" 9494..10229	
ORIGIN	
Query Match	
Best Local Similarity	83.7%
Score	6346.4
DB	9
Length	10229;
Matches	6625;
Conservative	194;
Mismatches	110;
Indels	52;
Gaps	28;
Db	3068
Qy	606
Db	2888
Qy	666
Db	2948
Qy	726
Db	3077
Db	3008
Qy	785
Db	3128
Qy	845
Db	3068
Qy	905
Db	3188
Qy	963
Db	3248
Qy	1024
Db	3367
Qy	1144
Db	3427
Qy	1184
Db	3487
Qy	1264
Db	3546
Qy	1324
Db	3603
Qy	1384
Db	3662
Qy	1444
Db	3719
Qy	1504
Db	3778
Qy	1564
Db	3837
Qy	1624
Db	3897
Qy	1682
Db	1683
Db	3957
Qy	1743
Db	4017
Qy	1803
Db	4077

Db	6262 TAGCAGACTAATACGCCCTACTCGGGCACTAGAAATTAGGAGAACAAAAGGSCAATA 6321	Qy
Qy	4083 TATATACAGCTTAACTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 4142	Db
Db	6322 TATATACAGCTTAACTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 6381	Qy
Qy	4143 AAAGGAAATCTACTCTCTGAGAACCTATGAGAACCTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 4202	Db
Db	6382 AAAGGAAATCTACTCTCTGAGAACCTATGAGAACCTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 6441	Qy
Qy	4203 TATATYTSACWGTTAGAGAACCTATGAGAACCTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 4262	Db
Db	6442 TATATYTSACWGTTAGAGAACCTATGAGAACCTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 6442	Qy
Qy	4263 AAAGGAAAGRAAGGGAAATASAGRGAATGCGAACGAACTATGAGGMAAAGAGCTG 4322	Db
Db	6502 AAGGAAAGRAAGGGAAATASAGRGAATGCGAACGAACTATGAGGMAAAGAGCTG 6561	Qy
Qy	4323 TCGGGAGAACCTCCATAGAGAACCTATGAGAACCTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 4382	Db
Db	6562 TCGGGAGAACCTCCATAGAGAACCTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 6621	Qy
Qy	4383 TCGGGAGAACCTCCATAGAGAACCTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 4441	Db
Db	6622 TCGGGAGAACCTCCATAGAGAACCTATGCTTAACTAGTCNTCATGCCATGMRGAATATGSARAG 6681	Qy
Qy	4442 AGTTTCTCCCTGGGAGGTAGGCACTGAGAGGGAAATCTTGTGCTGCACT 4501	Db
Db	6682 AGTTTCTCCCTGGGAGGTAGGCACTGAGAGGGAAATCTTGTGCTGCACT 6741	Qy
Qy	4502 ATCCATGAAATTCTAAACCCCTCAACACTTCACTTAGGCAACTGATGACCC 4561	Db
Db	6742 ATCCATGAAATTCTAAACCCCTCAACACTTCACTTAGGCAACTGATGACCC 6801	Qy
Qy	4562 CATCARATGCCAAATCATTTATCTGAGCAGGGCTTTCACAACTACTTCACCTTGACTAGGCATCGATGACCC 4621	Db
Db	6802 CATCARATGCCAAATCATTTATCTGAGCAGGGCTTTCACAACTACTTCACCTTGACTAGGCATCGATGACCC 6861	Qy
Qy	4622 KTCAGSCTCTGAAKGTGCCRARAATAATCCCTGCTTATCGCAAGCTCTCA 4681	Db
Db	6862 KTCAGSCTCTGAAKGTGCCRARAATAATCCCTGCTTATCGCAAGCTCTCA 6921	Qy
Qy	4682 GGCAGAACCTGAAAGAACCTGAACTCTTCAACACTTCACTTGACTAGGCATCGATGACCC 4741	Db
Db	6922 GGCAGAACCTGAAAGAACCTGAACTCTTCAACACTTCACTTGACTAGGCATCGATGACCC 6981	Qy
Qy	4742 CCAAACTCAGGATTCTAGTACTACTGTCTGGTARABACTTCACTGGTTGSCAR 4801	Db
Db	6982 CCAAACTCAGGATTCTAGTACTACTGTCTGGTARABACTTCACTGGTTGSCAR 7041	Qy
Qy	4802 AGGCCTTCCCCTGAGACAGAACGGCCAAAGGGTATAAGGCACTGTCTCA 4861	Db
Db	7042 AGGCCTTCCCCTGAGACAGAACGGCCAAAGGGTATAAGGCACTGTCTCA 7101	Qy
Qy	4862 TAATTCAGATTCGACTTCCCGAGCTTACAGAGAACCTGAACTGCTTCA 4921	Db
Db	7102 TAATTCAGATTCGACTTCCCGAGCTTACAGAGAACCTGAACTGCTTCA 7161	Qy
Qy	4922 CCACAGTAACTCCAGGAGTATCCCAAGCGTCTACAGAGAACCTGCTTCA 4981	Db
Db	7162 CCACAGTAACTCCAGGAGTATCCCAAGCGTCTACAGAGAACCTGCTTCA 7221	Qy
Qy	4982 GAGGGCACAGTCTCTGGGAGGTGAGAACCTGAGAACCTGCTTCA 5041	Db
Db	7222 GAGGGCACAGTCTCTGGGAGGTGAGAACCTGCTTCA 7281	Qy
Qy	5042 AAAGGAAACCGGAAACCCCTCACTGAGAACCTGCTTCA 5101	Db
Db	7282 AAAGGAAACCGGAAACCCCTCACTGAGAACCTGCTTCA 7341	Qy
Qy	5102 GAATCTGCAACTTCCCAAAGGGGACTTACGCATAGAAGTGTGAGG 5161	Db
Db	5162 CTTCATACCAATGACCTGTGTTGACCCAGAACGCAACTTAGTGAGACATAC 5221	Qy
Db	7402 CTTCATACCAATGACCTGTGTTGACCCAGAACGCAACTTAGTGAGACATAC 7460	Qy
Qy	5222 CTCTTACGCCAAATCATCACAGCTCTTAAACATTAACAGGAACTATCCCTGAGAAGA 5281	Db
Db	7461 CTCTTACGCCAAATCATCACAGCTCTTAAACATTAACAGGAACTATCCCTGAGAAGA 7520	Qy
Qy	5282 GGGAAAGRACTATTCACCCWMTGACATGCTTACGTTAGTCAGTCCCTCYCTAATC 5341	Db
Db	7521 GGGAAAGRACTATTCACCCWMTGACATGCTTACGTTAGTCAGTCCCTCYCTAATC 7580	Qy
Qy	5342 CCCATCCTAGATCATCTGGAGAACCTCCAGTCATTATACCCAACTGC 5401	Db
Db	7581 CCCATCCTAGATCATCTGGAGAACCTCCAGTCATTATACCCAACTGC 7640	Qy
Qy	5402 GTTAATAGTGGCTGGAGTGTGAGTCTGCTGATACATCACCTGAGTCATACT 5461	Db
Db	7641 GTTAATAGTGGCTGGAGTGTGAGTCTGCTGATACATCACCTGAGTCATACT 7700	Qy
Qy	5462 GCGAAAGRACTGAAATCCGGAGAACCTGACTGTTGAGTACATCACCTGAGTCATACT 5521	Db
Db	7701 GCGAAAGRACTGAAATCCGGAGAACCTGACTGTTGAGTACATCACCTGAGTCATACT 7760	Qy
Qy	5522 TTGCGCTGCTCTAACACACAGGAGGAAGTACTTAATCCCAACTGC 5581	Db
Db	7761 TTGCGCTGCTCTAACACACAGGAGGAAGTACTTAATCCCAACTGC 7819	Qy
Qy	5582 TGGSCCTCTCATATCATATTTCCTCTTCTCTTCTCTTCTCTTCTCTCTACT 5641	Db
Db	7820 TGGSCCTCTCATATCATATTTCCTCTTCTCTTCTCTCTACT 7879	Qy
Qy	5642 CCCCCCTCAGGCCSTGATGAGCGTAGSTCCCTTYACCCSTTTCACCTCTACTG 5701	Db
Db	7880 CCCCCCTCAGGCCSTGATGAGCGTAGSTCCCTTYACCCSTTTCACCTCTACTG 7939	Qy
Qy	5702 TCGAGCTCCGAAATATGNGCCATGCTTCTAGGAGCTTCTAGGAGAACCCCA 5761	Db
Db	7940 TCGAGCTCCGAAATATGNGCCATGCTTCTAGGAGAACCCCA 7999	Qy
Qy	5762 CCTTCACTGCCAACCCATTTGCCCCGCAACTGCTTACACTCTGCACTCTTGCTG 5821	Db
Db	8000 CCTTCACTGCCAACCCATTTGCCCCGCAACTGCTTACACTCTGCACTCTTGCTG 8059	Qy
Qy	5822 ATGCAAATACTATTATGGAGGAAATGATTAATCTCTGTTGCTGCTGAGGACTTG 5881	Db
Db	8060 ATGCAAATACTATTATGGAGGAAATGATTAATCTCTGTTGCTGCTGAGGACTTG 8119	Qy
Qy	5882 GGTCACTGTGCTGAGACTACTGACCAACTGGTATGCTGATGGGGTGGAGTC 5941	Db
Db	8120 GGTCACTGTGCTGAGACTACTGACCAACTGGTATGCTGATGGGGTGGAGTC 8179	Qy
Qy	5942 AGATCAAGGAGGAAATGTTAAAGGAGTAATCTCCAACTCACCSCGGTACATG 6001	Db
Db	8180 AGATCAAGGAGGAAATGTTAAAGGAGTAATCTCCAACTCACCSCGGTACATG 8239	Qy
Qy	6002 GCACCTCTASCCCTACAAGGAGTAATCTCAAAACTGATGAACTGCTTCA 6061	Db
Db	8240 GCACCTCTASCCCTACAAGGAGTAATCTCAAAACTGATGAACTGCTTCA 8299	Qy
Qy	6062 ATACTGCGCTGCTGAGCTTAACTACCCCTACTGGCTCATGGCTCTGCC 6121	Db
Db	8300 ATACTGCGCTGCTGAGCTTAACTACCCCTACTGGCTCATGGCTCTGCC 8359	Qy
Qy	6122 AACCTACTACTGTGGATATGCTTCCCTGACTCTGCACTGCTTCA 6181	Db
Db	8360 AACCTACTACTGTGGATATGCTTCCCTGACTCTGCACTGCTTCA 8419	Qy
Qy	6182 CTGACTGAGAACCTGAGAACCTGAGAACCTGAGAACCTGAGAACCTGAG 6241	Db
Db	8420 CTGACTGAGAACCTGAGAACCTGAGAACCTGAGAACCTGAGAACCTGAG 8479	Qy

QY 6242 GACCTCTGTTCATAATSGAAATAACCAACTTCAAACCTCACCTGTTAATT 6301
Db 8480 GACCTCTGTTCACAATCTGGAAATAACCAACTTCAAACCTCACCTGTTAATT 8539
QY 6302 GCAATCTACTACAGACCAACTCCATCATCAGGTGTAACCTCCACACAA 6361
Db 8540 GCAATCTACTACAGACCAACTCCATCATCAGGTGTAACCTCCACACAA 8599
QY 6362 TAGTCGCCTAACCTCGAGAATATTGTTGTTGTTACTCGCTATGGTTGA 6421
8600 TAGTCGCCTAACCTCGAGAATATTGTTGTTGTTACTCGCTATGGTTGA 8659
Db 6422 ATGGCTCTTCGAATCTATGCTCTCTATCTAGTGGCCCTATGCCATGTTGA 6481
QY 8660 ATGGCTCTTCGAATCTATGCTCTCTATCTAGTGGCCCTATGCCATGTTGA 8719
Db 6482 CTGAACTAGATTAACAGTAGTCATAATTAGCCGACAAAGAGTACCCATC 6541
QY 8720 CTGAACTAGATTAACAGTAGTCATAATTAGCCGACAAAGAGTACCCATC 8779
QY 6542 TCCCTTGTTAGGAGCAGGAGTGGCTAGGGACTGGATTCGGGATCA 6601
Db 8780 TCCCTTGTTAGGAGCAGGAGTGGCTAGGGACTGGATTCGGGATCA 8839
QY 6602 GACCTCTACTCAGTCTACTCAAACTATCAGAGCTATGGGACTGGAACGG 6661
Db 8840 GACCTCTACTCAGTCTACTCAAACTATCAGAGCTATGGGACTGGAACGG 8899
QY 6662 TGGCCACTCTGGTACCTGCAACTTAACCTCTGGTACCTGAGCTGGCTTC 6721
8900 TGGCCACTCTGGTACCTGAGCTGGCTACATCTGGTACCTGAGCTGGCTTC 8959
QY 6722 AAATCGAGAGCTTAGACTTGCTAACCGCTGAGAGGGAACTGTTATTAG 6781
Db 8960 AAATCGAGAGCTTAGACTTGCTAACCGCTGAGAGGGAACTGTTATTAG 9019
QY 6782 GGGAGAGATGCGTTATGTTATCATCCGGAATGTTCTGAGAAAGTAAAGAA 6841
9020 GGAGAGATGCGTTATGTTATCATCCGGAATGTTCTGAGAAAGTAAAGAA 9079
QY 6842 TCCSAGATGAACTACAAGTAGCAGARGACTTGGACCTGGGCTTC 6901
Db 9080 TCCGAGATGAACTACAAGTAGCAGAGGAGCTGGGCTTC 9139
QY 6902 TCAAGGCGATGATGCCCTGATTCCTCTTAGGCTCTAGCCTATATG 6961
Db 9140 TCAAGGCGATGATGCCCTGATTCCTCTAGGCTCTAGCCTATATG 9199
QY 6962 TACTCCCTTGGACCCCTGATTCCTCTAGGCTCTAGCCTATATG 7021
Db 9200 TACTCCCTTGGACCCCTGATTCCTCTAGGCTCTAGCCTATATG 9259
QY 7022 AAGCTGTTAACTACAATGGACGCCAGATGCGTCCAGACTAGATCTCCGAGAC 7081
Db 9260 AAGCTGTTAACTACAATGGACGCCAGATGCGTCCAGACTAGATCTCCGAGAC 9319
QY 7082 CCTGGCCGCGCTGTYGACGCCAGATGTTAATGCACTCAAGGCCCTCTG 7141
Db 9320 CCTGGCCGCGCTGTYGACGCCAGATGTTAATGCACTCAAGGCCCTCTG 9379
QY 7142 AGGAATCTCAGTCAGACCTACTAGGCCAATTCAGGAGGAGCTAGGCC 7201
Db 9380 AGGAATCTCAGTCAGACCTACTAGGCCAATTCAGGAGGAGCTAGGCC 9439
QY 7202 GTSGTCGCCAACCTCCCACACAGCACTTAGGTTCTGGTGGAGTGGAG 7261
Db 9440 GTGTCGCCAACCTCCCACACAGCACTTAGGTTCTGGTGGAGTGGAG 9499
QY 7262 ACUGGACTAGCTGATTCCTAGCTGAGATCTCYTAASCTCTAGTGGAGGAG 7321
Db 9500 ACAGGACTAGCTGATTCCTAGCTGAGATCTCYTAASCTCTAGTGGAGGAG 9559

QY 7322 CCACATCACCCTTAAACGGGCTTCAACTTAGTCACCTGACGAC 7381
Db 9560 CCACATCACCCTTAAACGGGCTTCAACTTAGTCACCTGACGAC 9619
QY 7382 TCACTAACTGCTTAAAGCAGAGGAGTAAAGAAATAGCCATATG 7441
Db 9620 TCACTAACTGCTTAAAGCAGAGGAGTAAAGAAATAGCCATATG 9679
QY 7442 TGAGACAGCAGGAGGAGTAAAGCAGAGGAGTAAAGAAATAGCCATATG 7501
Db 9680 TGAGACAGCAGGAGGAGTAAAGCAGAGGAGTAAAGAAATAGCCATATG 9739
QY 7502 GGCAACCCCTTGGTCCCTCTGTTATGGAGCTGTTCTATCTTACT 7561
Db 9740 GGCACCCCTTGGTCCCTCTGTTATGGAGCTGCTGTTCTATCTACT 9799
QY 7562 CTATTAACTTGCACTG 7582
Db 9800 CTATTAACTTGCAACTG 9820

RESULT 15

AY101587 Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVWEI

DEFINITION Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVWEI

LOCUS locus, allele B, complete sequence.

ACCESION AY101587

VERSION AY101587.1 GI:37544469

KEYWORDS

- Pan troglodytes (chimpanzee)
- Pan troglodytes
- Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarhini; Hominidae; Pan.

REFERENCE (bases 1 to 10229)

AUTHORS Malli, F., Boulon, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonneau, B., Ducote, G., Duret, L. and Mandrand, B.

TITLE The endogenous retroviral locus ERVWEI is a bona fide gene involved in hominid placental physiology

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)

PUBMED 14757826

REFERENCE

AUTHORS Mallet, F., Boulon, O. and Oriol, G.

TITLE Direct Submission

JOURNAL Submitted (05-MAY-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'italie, Lyon 69364 cedex 07, France

FEATURES

Source

1. 10229 /organism="Pan troglodytes"
- /mol_type="genomic DNA"
- /isolate="1"
- /isolation_source="PBMC of chimpanzee 1"
- /db_xref="Taxon:958"
- /cell_type="PBMC"
- /note="isolated by PCR; ERVWEI locus, allele B endogenous virus: ERV-W"
1. .779
- Join(248, .810-.9820)
- /note="putative mRNA transcript 1; gag-pol pseudogene"
- /join(248, .810-.9820)
- /product="envelope glycoprotein"
- /note="putative mRNA transcript 2"
- /join(248, .810-.9820)
- /note="putative mRNA transcript 3; no coding region"
- 7819 .9435
- /note="syncytin"
- /codon_start=1
- /product="envelope glycoprotein"
- /protein_id="AAW68166.1"
- /db_xref="GI:37544469"
- /translation="MAlPwIFLFTVLPSFLTAPPCCRWTSSSPYQDFLWRMOPGNIDPSPYSPSKGPPTFTTHMRNCYSATLCMHNHYWKGKINNSCPGGGIVTVCWTFIHGMSDGGVQDQAREKVKVEVSQLTRVHSASRPYKQSLDKLHETLT

HPLVLSLNTLTLHEVSQAQPNCILPANRPRVSPVPEQMANSTINTSV
 LYCPPLVNLSELETTITNLTICVKSENNTDTINSQCRWTPPTOIVCPSGIFTWVGNSA
 TCGGTTTSTOFYKUOBINGMERADSLTQLDQNLNSLAUAVLQNRRAIDLTAE
 RGCICLIGECCYVNGSIVEKEVIRDIORRERLNGTGPMILSOMPWLP
 FPLGLAATILLIFGPCLPNLNUFNEASRTEAVKLQMEPKVOSKTKIYRRPDRPASP
 RSDVNDKGTPPEBLTQPLRPNSAGS" 9494. . 10229

ORIGIN

LTR	Query	Match	Start	End	Score	Length	DB	Best Local Similarity	Conservative	Mismatches	Indels	Gaps
	Qy	606	1504	3897	CAGGGGAGGAGAAATCTRACTGCTTCTGGAGAGACTAAGGGAGCATGGGA 83.7%	10229	Db	154	1563	111	52	28;
	Db	2888	3778	CA-GGGAGGGAGAAATCTRACTGCTTCTGGAGAGACTAAGGGAGCATGGGA 655	2888	2948	GATGGAAAAGCTCCCGAGAACGACAAAGGCCCTAAGCAGATTGARATTGGA 2947	666	665	111	52	28;
	Qy	726	1564	3837	GCCTCTGGAGGAGATAATATAACACCTCTTACAGTAGCTAGACTCTTGTAGA 1623	10229	Db	1564	1563	111	52	28;
	Db	3008	3837	GCCTCTGGAGGAGATAATATAACACCTCTTACAGTAGCTAGACTCTTGTAGA 3067	3008	3068	GCACCTCTGAGGGAGATAATATAACACCTCTTACAGTAGCTAGACTCTTGTAGA 3067	786	785	111	52	28;
	Qy	846	1624	4137	CACTCACTCAGCTGAGCTAG-AAAAACTTCAGAACGCTGCCCAGGCCGGCCA 1682	10229	Db	1624	1623	111	52	28;
	Db	3128	3897	CACTCACTCAGCTGAGCTAG-AAAAACTTCAGAACGCTGCCCAGGCCGGCCA 3956	3128	3128	AATTATGTAAGAAGTGTATTATGCCCTACAGGAAGCCTACTGCTACTCCATC 3187	906	905	111	52	28;
	Qy	964	1803	3957	AAACTTGAACCCATTGAACTTGTGCTACTGGCTACTCCTTATATAGAGATCAGGGAA 3067	10229	Db	1803	1803	111	52	28;
	Db	3248	4016	GCAGCTGAGCTTGGAGGACTTGTGAAAGGAAAGCTGGGCAATGTAAAGCTATA 1862	3248	3248	AAAGGGAGATGACAAAGGGTAACAGTGACCAAGGCTAACAGTACACTCTTGTAGA 3247	964	963	111	52	28;
	Qy	1024	4076	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 1922	1024	1024	GACCCCTCCCTGACTCCTCCAACTTAAGGACCCCTCACCAAATGTC 3247	964	963	111	52	28;
	Db	3308	4136	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4196	1024	3308	GACCCCTTCCAGCAGCTGGAGGAAGKAATGTCAGTAGAGTA 3366	964	963	111	52	28;
	Qy	1084	4136	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4196	1024	1084	TTCTCTCAGCTTAAGCAATAAACACACTTAGGTAATTCCAGATACCTG 3307	964	963	111	52	28;
	Db	3367	4256	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4256	1024	3367	TCTCTCTCAGCTTAAGCAATAAACACACTTAGGTAATTCCAGATACCTG 3426	964	963	111	52	28;
	Qy	1144	4256	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4256	1024	1144	GGCTATATGTTTACAGGAGTACCTCTTGTGATCATGGAGAGATA 1143	964	963	111	52	28;
	Db	3427	4316	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4316	1024	3427	GGCTATATGTTTACAGGAGTACCTCTTGTGATCATGGAGAGATA 1143	964	963	111	52	28;
	Qy	1204	4316	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4316	1024	1204	TATGTCACTGCTAATCAGACACTAACCCCAATGAGAGATGCCACATCTGC 1263	964	963	111	52	28;
	Db	3487	4355	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4355	1024	3487	-ATGTCACTGCTAATCAGACACTAACCCCAATGAGAGATGCCACATCTGC 3487	964	963	111	52	28;
	Qy	1264	4355	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4355	1024	1264	CTGAGRTGTTGGCATCTCTGTTACAGTATGAGTATGAGTATGAGTACCTG 1263	964	963	111	52	28;
	Db	3546	4436	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4436	1024	3546	CTGAGRTGTTGGCATCTCTGTTACAGTATGAGTATGAGTACCTG 1263	964	963	111	52	28;
	Qy	1324	4436	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4436	1024	1324	CTGAGRTGTTGGCATCTCTGTTACAGTATGAGTATGAGTACCTG 1263	964	963	111	52	28;
	Db	3603	4551	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4551	1024	3603	GGAGAGAGAGATGTTCCCGAGGAGCAGTCAGCTTGTAGACCTCATGGG 3661	964	963	111	52	28;
	Qy	1384	4551	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4551	1024	1384	ACACAGAAATCAGTACATGGAGATGGCTGAGCATTTCTACTTGTGCTA 1443	964	963	111	52	28;
	Db	3662	4670	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4670	1024	3662	-ACACAGAAATCAGACAT-GGAGATGGCTGAGCATTTCTACTTGTGCTA 3718	964	963	111	52	28;
	Qy	1444	4670	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4670	1024	1444	AGGAGCTAAGGAAACTASAGAGAARTCTAGAATTACATGTCACCTA 1503	964	963	111	52	28;
	Db	3719	4769	GGGGAGAAAGCTTGTGAGGAGCTTAAAGAATGTCAGTAGAGTA 4769	1024	3719	AGGAGCTAAGGAAACTAGGAG-AAGCTTATGAACTCTAACATGTCACCTA 3777	964	963	111	52	28;

QY	2583	GCCCCAGGCTTAGTAAACCATGCTTAACCTCCCTGCAGTAATCCGTAGTGATTGAGGA	2642	Db	5903	GAATGGCATACTCACTGTAAAGGAGACTTGTGACTGTGAGAACACTTTTAAATG	5962
Db	4850	GCCCCAGGCTTAGTAAACCATGCTTAACCTCCCTGCAGTAATCCGTAGTGATTGAGGA	4904	QY	3723	TGCGCTTATIACCTGARGGCCAGTCTGCRACCTGTCACCTCTAACC	3782
QY	2643	GGCAAGAAACCCAGGAGCAGTGAGGGTTAGTGCAGATCTCAGATTCAATGG	2702	Db	5963	TCAAGCTTAACTGTGAGACGGCTGCTGCACCTGCACTCTAACCC	6022
Db	4905	ATTAAAGAACCCACAGACAGTGGA-GGTAGTGTGCAAGATCTCAGATTCAATGAG	4962	QY	3783	AGCNCATTTCTCCAGACATAGAGAAAGATAAAATACGTCAAAARATTTC	3842
QY	2703	GCCGGTGCCTTTACCCAGCTGACCTACCCCTATACGTGMYTICCAATCCA	2762	Db	6023	AGGCCACATTCTCCAGACATAGAGAAAGATAAAATACGTCAAAARATTTC	6082
Db	4963	GCTGTGTTCTCTATGGCAGCTGACTGACCTTACCTGCTTCCAAATCCA	5022	QY	3843	TCAACCTPATGCACTCGAGGSCACCTTGTAGGTTGACTGATCCTCA	3902
QY	2763	GAGGAGCAGAGTGTITACASTCTCTGACCTTACGGATGCTTCTTGCACTCCCTGTA	2822	Db	6083	TCAACCTPATGCACTCGAGGSCACCTTGTAGGTTGACTGATCCTCA	6141
Db	5023	CATCTGACTCTCAATCTGTTGCCTTGAGAATCTCAAACCCARCATCTCAACT	5082	QY	3903	ACTGTATACTGTGAGAATTCTGAGAAAGAGCTCGAAAGGGGATATGAG	3962
QY	2823	CATCTGACTCTCAATCTGTTGCCTTGAGAATCTCAAACCCARCATCTCAACT	2882	Db	6142	ACTGTATACTGTGAGAATTCTGAGAAAGAGCTTGAAAGGGGATATGAG	6201
Db	5083	CATCTGACTCTCAATCTGTTGCCTTGAGAATCTCAAACCCARCATCTCAACT	5142	QY	3963	TGTCAGCTGATAATGGATAATTGAGGATGYCCCATCTATTGGCAGGCATTA	4022
QY	2883	ACCTGACTRTTACCCAAAGGGTCAGGGATAGYCCCATCTATTGGCAGGCATTA	2942	Db	6202	TGGTCAGGGATAATGGATAATTGAGGATGYCCCATCTATTGGCAGGCATTA	6261
Db	5143	ACCTGACTCTTACCCAAAGGGTCAGGGATAGYCCCATCTATTGGCAGGCATTA	5202	QY	4023	TRGCAGACTTAATGGCTTGTAGGAACTTGTGCTCTCGTAKTGAGAT	4082
QY	2943	GCCCCAGACTTGAGKARTYMCATCTGGCACACTGTCGCTCATCTGGACATCT	3002	Db	6262	TACAGAGACTTAATGGCTTGTAGGAACTTGTGCTCTCGTAKTGAGAT	6321
Db	5203	GGCCAGACTTGAGCCATCTGGACATACTGGACATCTGGCTTGTGCTCTCGTAKTGAGAT	5262	QY	4083	TATATACAGACTCTRATAATGGCTTGTAGGAACTTGTGCTCTCGTAKTGAGAT	6381
QY	3003	TTACTTTRGCGCCATTCAGAACCTTGCCATAGGCCACCAAGCAGCTTMAAT	3062	Db	6322	TATATACAGACTCTRATAATGGCTTGTAGGAACTTGTGCTCTCGTAKTGAGAT	6322
Db	5263	TTACTTTRGCGCCATTCAGAACCTTGCCATAGGCCACCAAGCAGCTTMAAT	5322	QY	4143	AAAGGGATTCCTRACTHYGAGGAACTCATAMATCAGGAGCAATTAGGAA	4202
QY	3063	TTCCCTGGCYACCTGTGCTACAWGGTTCCAAOSARAGTCARCTGTCACAGAG	3122	Db	6382	AAAGGGATTCCTRACTCTGAGGAACTCATAGGAACTTACCTGAGGAACTTAC	6441
Db	5323	TTCCCTGGCTTACCTGTGCTACATGTGTTCAACAAAGGTCACCTGTCACGGCG	5382	QY	4203	TATTAATGGCTACAGAACCTTGTGAGGATGGAGTCTACCTGTCACGGCTC	4262
QY	3123	GT---TACTTAGGCTAAATTATCCAAAGGACCCAGGGACCTCTAGTGAGGACAC	3182	Db	6442	TATTATGGCTACAGAACCTTGTGAGGATGGAGTCTACCTGTCACGGCTC	6501
Db	5383	GT---TACTTAGGCTAAATTATCCAAAGGACCCAGGGACCTCTAGTGAGGACAC	5438	QY	4263	AAGGAAGRAAGGAAATASAGRGAAVTGCCAGCATATTGAGCAGAACGAT	4322
QY	3183	CAGCTCTACTGGCTTATCTCATCYAAACCCATAAGCAGCTAAGRGRRTCTCTGG	3242	Db	6502	AAGGAAGGAAGGAAATAGAGAAGACTGCCCAGCAGATAGGAAAGGAG	6561
Db	5439	CAGCTCTACTGGCTTATCTCATCCAAACCTTAAGGACACTAAGGAGATCTCTGG	5498	QY	4323	CAAGGAGGAAACCCCTCATTAAGAATGCTTAACTCTCCCTAGTAAAGGAA	4382
QY	3243	RTAYAGGYTCTGCGGAAWATGGATTCCAGGTWTGGCRAATAGCCAGGYCATTAA	3302	Db	6562	CAAGGAGGAAACCCCTCATTAAGAATGCTTAAACCAACCCCTAGTAAAGGAA	6621
Db	5499	GTPATAGTTCTGCCAAATATGGATT-CCCGGTATGGCAAAATGGCAGGCTTAA	5557	QY	4383	TCCGGAAACCAAGCCCCTGACTCAGCGGGAAACAGAATGGGAACTCTACGGG-C	4441
QY	3303	TACATTTAGGAACTCAGAACCTTACCCATTARTAGATGAGTGAATGAGTGAAT	3362	Db	6622	AGTTTCTCCCTGGGATGCTGAGCACTGAGGAAATACTTCTGCAACT	6681
Db	5558	TACATTTAGGAACTCAGAACCTTACCCATTAGTGAATGAGTGAATGAGTGAAT	5617	QY	4442	AGTTTCTCCCTGGGATGCTGAGCACTGAGGAAATACTTCTGCAACT	6741
QY	3363	MRAAGTGCTTCCAGGCCCTAACCCAGGCTTAACCCAGGCCAGTGTAGYTG	3422	Db	6682	AGTTTCTCCCTGGGATGCTGAGCACTGAGGAAATACTTCTGCAACT	6741
Db	5618	AGAAAGGCTTCCAGGCCCTAACCCAGGCTTAACCCAGGCCAGTGTAGYTG	5663	QY	4502	ATCCATGGAATRACTTAAACCTCTCAACCTTCACTGGCATGATAGAAC	4561
QY	3423	CCAAACRGGGCAAGACTTTSTTYATYTCACAGAAACAGGAAVAGCTCTGGAGT	3482	Db	6742	AACCATGGAATRACTTAAACCTCTCAACCTTCACTGGCATGATAGAAC	6801
Db	5664	CCAAACGGGCAAGACTTTCTCTCATGTCAGCAG-AAAACAGGAATGGCTTAA	5722	QY	4562	CATCARATGCCAAATCATATTACTGAGGACCTGAGGAAATACCTGCAACT	4621
QY	3483	CCTTACACAGATCAGGGATGACTGCACTGAGGATACCTGACTGAGGAA	3542	Db	6802	CATCARATGCCAAATCATATTACTGAGGACCTGAGGAAATACCTGCAACT	6861
Db	5723	CCTTACACAGATCAGGGATGACTGCACTGAGGATACCTGACTGAGGAA	5782	QY	6822	KTCAGGGCTCTGAGTGTGCACTGAGGAAATACCTGCAACT	6981
QY	3543	TGAGTGGCAAGGGTTCRGCYTCATGTGTTAGTGGCTGAGCTAGCAGTGTAG	3602	Db	6862	GTCAAGGGCTCTGAGTGTGCAACT	6921
Db	5783	TGAGTGGCAAGGGTTCRGCYTCATGTGTTAGTGGCTGAGCTAGCAGTGTAG	5842	QY	6862	GGARACAAACAGGGCCATTACCTGAGGAAATACCTGCAACT	7001
QY	3603	ATCTGAGGCACTACTCTGTAATAGGGAGATGTGCTGAGCTACCTGAGT	3662	Db	6922	GGAGACAAACAGGGCCATTACCTGAGGAAATACCTGCAACT	7001
Db	5843	ATCTGAGGCACTACTCTGTAATAGGGAGATGTGCTGAGCTACCTGAGT	5902	QY	6922	GGAGACAAACAGGGCCATTACCTGAGGAAATACCTGCAACT	7001
QY	3663	GAAYGCACTACTCTGTAATAGGGAGATGTGCTGAGCTACCTGAGT	3722	Db	6942	CCAAACCTGGGAAATCTGAGGAAATACCTGCAACT	7001

Db	6 982	CCAAATCTCGGGATTCAAGTACTRACTACTCTGGTAGTACTTCAGGGGGAGG 7041	QY	5 882	GACGTCACTCTCTGACTTACTTCACCAAACTGTTGTTGTTGATGGGGGGAGGTC 5941	QY
QY	4 802	AGGCCTTCCCCTGTAGGAGCAGRAAAGGCCAGGGTAATAAGGCACTGTTGAA 4861	Db	8 120	GAGTCAGTCGTTGACTTACTTCACCATCTGGTATGCTGATGGGGGGAGGTC 8179	Db
Db	7 042	AGGCTTCCCCTGTAGGAGCAGRAAAGGCCAGGGTAATAAGGCACTGTTGAA 7101	QY	5 942	AAGATCGGCAAGAGAAGAACATGTAAGAAGTAATTCGAACACTACCCSGGGTACAGC 5001	QY
QY	4 852	TAATTCAGGATGGACTTCCCGAGGTTACAGAGTGCAATAGCCCTGTTCCAGG 4921	Db	8 180	AAGATCGGCAAGAGAAGAACATGTAAGAAGTAATTCGAACACTACCCSGGGTACAGC 8239	Db
Db	7 102	TAATTCAGGATGGACTTCCCGAGGTTACAGAGTGCAATAGCCCTGTTCCAGG 7161	QY	6 002	GCACCTTGACCCCTACAAAGGACTTACATCTCAAAACTACATGAAACCTTGGTACCC 6061	QY
QY	4 922	CCACAGTAAACCCAGGGATATCCAGGGCTAGGATATACCAATCACTTACCTGCGCT 4981	Db	8 240	GCCCTTGTAGGCTTACAGAGACTTACATCTCAAAACTACATGAAACCTTGGTACCC 8299	Db
Db	7 162	CCACAGTAAACCCAGGGATATCCAGGGCTAGGATATACCAATCACTTACCTGCGCT 7221	QY	6 062	AATACCTCCTGTTAGCTTAACTTACCCACTTACGGCTTACGGCTTCAGGTCGGCC 6121	QY
QY	4 982	GAAGGCCAGTCTCAGGGAGGTGAGGAGTCAAGGAGTCAAGAACTCAAGGGCATCAA 5041	Db	8 300	AATACCTCCTGTTAGCTTAACTTACCCACTTACGGCTTACGGCTTCAGGTCGGCC 8359	Db
Db	7 222	GAAGGCCAGTCTCAGGGAGGTGAGGAGTCAAGGAGTCAAGAACTCAAGGGCATCAA 7281	QY	6 122	AAACCTACTACTGTGGATATGCCCTGAACTCARGCCATATGTTCAATCC 6181	QY
QY	5 042	AAAAGCAACCCAGGAACCACCTCACATGGCTGTYCTGTTGCTATAGCTTAAGA 5101	Db	8 360	AAACCTACTACTGTGGATATGCCCTGAACTCARGCCATATGTTCAATCC 8419	Db
Db	7 282	AAAAGCAACCCAGGAACCACCTCACATGGCTGTYCTGTTGCTATAGCTTAAGA 7341	QY	6 182	CITGACTCTGACGATGGACACTCGACAGCCAACTTAGTGTGAGACATCAC 6241	QY
QY	5 102	GAATCTGCACTTCCCAAAGCAGGACTTACCCATACTGAATGTTGAGAACGC 5161	Db	8 420	CITGACTCTGACGATGGACACTCGACAGCCAACTTAGTGTGAGACATCAC 8479	Db
Db	7 342	GAATCTGCACTTCCCAAAGCAGGACTTACCCATACTGAATGTTGAGAACGC 7401	QY	6 242	GACCTCTGTTCCAACTCTGGATATTAACCCATACCCACTTACCGTTGAGG 6301	QY
QY	5 162	CCITCATAACCAACTTCCCAAAGCAGGACTTACCCATACTGAATGTTGAGAACGC 5221	Db	8 480	GACCTCTGTTCCAACTCTGGATATTAACCCATACCCACTTACCGTTGAGG 8539	Db
Db	7 402	CCITCATAACCAACTTACGACTTGTGCTG-CAGACAGCCAACTTAGTGTGAGACATCAC 7460	QY	6 312	GCAATCTACTACATACACCAACTCCAACTCACTGGCTGAACTTCCACCAA 6361	QY
QY	5 222	CCCTCTAGCCAAATATCACAACTTCAACATACAGGACCATCTGGAGAGA 5281	Db	8 540	GCAATCTACTACATACACCAACTCCAACTTCACGCTTCAACCAA 8599	Db
Db	7 461	CCCTCTAGCCAAATATCACAACTTCAACATACAGGACCATCTGGAGAGA 7520	QY	6 362	TAGTCCTTACCTCTAGGAATATTTCCTGTTGACTTCAGGCTTCAATTC 6421	QY
QY	5 282	CCGAAGAACTATCCGCCWGGTACATGGTATAGTCAGTCAGTCTCTCTCTAATC 5341	Db	8 600	TAGTCCTTACCTCTAGGAATATTTCCTGTTGACTTCAGGCTTCAATTC 8659	Db
Db	7 521	CCGAAGAACTATCCGCCWGGTACATGGTATAGTCAGTCAGTCTCTCTAATC 7580	QY	6 422	ATGGCTTCTGTTCCAATSTGGAATTAACCCATACCCAACTCTGTTGTAATTA 6481	QY
QY	5 342	CCCATCCCTGATACATCTCTGGAGGACCTACCCAGCTTATTTACCCAACTGC 5401	Db	8 660	ATGGCTTCTGTTCCAATCTGGATATACCCATACCCAACTTCAGGCTTCAAC 8719	Db
Db	7 581	CCCATCCCTGATACATCTCTGGAGGACCTACCCAGCTTATTTACCCAACTGC 7640	QY	6 482	CTGAACAGATAATTAGTGTATCTGATTAATCTAAGCCCCAACAAAGTACCCATC 6541	QY
QY	5 402	GCTTAAGTGTGTTGGACTTCTGGATACATCACACTTCAGTCAACTCTGGTACT 5461	Db	8 720	CTGAACAGATAATTAGTGTATCTGATTAATCTAAGCCCCAACAAAGTACCCATC 8779	Db
Db	7 641	GCTTAAGTGTGTTGGACTTCTGGATACATCACACTTCAGTCAACTCTGGTACT 7700	QY	6 542	TCCTTGTATAGGACGAGGAGCTTCTGGTACTTCTGTTGAGCTTCTGTTGAGGA 6601	QY
QY	5 462	CCCAAGAAACTGAAATCAGGACAACCTAGTATTCCTGTTGAACTCTGAGGA 5521	Db	8 780	TCCTTGTATAGGACGAGGAGCTTCTGGTACTTCTGTTGAGCTTCTGTTGAGGA 8839	Db
Db	7 701	CCCAAGAAACTGAAATCAGGACAACCTAGTATTCCTGTTGAACTCTGAGGA 7760	QY	6 602	CAACCTCTACTGTTGACTTACCAACTATCTCAAGAACTTAATGGGACATGGAGGG 6661	QY
QY	5 522	TTGCGCCCTGTTCTCAACACACACCGGGAAGTAATAATCATAACTCCCCA 5581	Db	8 840	CAACCTCTACTGTTGACTTACCAACTATCTCAAGAACTTAATGGGACATGGAGGG 8899	Db
Db	7 761	TTGCGCCCTGTTCTCAACACACACCGGGAAGTAATAATCATAACTCCCCA 7819	QY	6 662	TCGGGATCTCTCTGTTCTGAACTTAATGGGACATGGAGCTTCTGGTACTTCTC 6721	QY
QY	5 582	TCGSCCTCTTATCATATTTCCTCTGTTCTGAACTTAATGGGACATGGAGCTTCTC 5641	Db	8 900	TCGGGATCTCTCTGTTCTGAACTTAATGGGACATGGAGCTTCTGGTACTTCTC 8959	Db
Db	7 820	TCGSCCTCTTATCATATTTCCTCTGTTCTGAACTTAATGGGACATGGAGCTTCTC 7879	QY	6 722	RAATCGAGAGCTTGTAGCTGAACTTAATGGGACATGGAGCTTCTGGTACTTCTC 6781	QY
QY	5 642	CAACCTCTCTGTTCTGAACTTAATGGGACATGGAGCTTCTGGTACTTCTC 5701	Db	8 960	AAATCGAGAGCTTGTAGCTGAACTTAATGGGACATGGAGCTTCTGGTACTTCTC 9019	Db
Db	7 880	CAACCTCTCTGTTCTGAACTTAATGGGACATGGAGCTTCTGGTACTTCTC 7939	QY	6 782	GGGAAGATACTGTTATAGTGTATCTCAATCCGGATCTGCACTGAGAATGTTAGGAA 6841	QY
QY	5 702	TGCAAGGTCCGGATATGGAGCTATGTTCTTAACTGGGACCTCCACCTCTGCGCT 5761	Db	9 020	GGGAGAACTGTTATAGTGTATCTCAATCCGGATCTGCACTGAGAATGTTAGGAA 9079	Db
Db	7 940	TGCAAGGTCCGGATATGGAGCTATGTTCTTAACTGGGACCTCCACCTCTGCGCT 7999	QY	6 842	TTCAGATGAACTGAACTTCTGGTACTTCTGGTACTTCTGGTACTTCTGGTACTTCTC 6901	QY
QY	5 762	CCTCACTGCGCACCCATGCGCCACTGCTATCACTCTGGTACTTCTGGTACTTCTC 5821	Db	9 080	TTGAGATGTTGACTTACAGTGTAGGAGCTTCTGGGCTTC 9139	Db
Db	8 000	CCTCACTGCGCACCCATGCGCCACTGCTATCACTCTGGTACTTCTGGGCTTC 8059	QY	6 902	TCACCGTGTGATCCCTGTTAGGACCTCTGGGCTTC 9199	QY
QY	5 822	ATGCAAACTCATATTGGAGGAAATTAATCTTACCTGTTGTCCTGGAGACTG 5881	Db	9 140	TCAGCCATGGATCCCTGTTAGGACCTCTGGGCTTC 9199	Db
Db	8 060	ATGCAAACTCATATTGGAGGAAATTAATCTTACCTGTTGTCCTGGAGACTG 8119				

us-10-717-580-11.rge

Search completed: November 16, 2005, 16:57:00
Job time : 31274 secs

Page
1

Copyright (c) 1993 - 2005 Compugen Ltd.
Om nucleic - nucleic search, using sw model
Run on: November 15, 2005, 21:00:58 ; Search time 1110 Seconds
(without alignments)
 11176.802 Million cell updates/sec
Title: US-10-717-580-11
Perfect score: 7582
Sequence: 1 caaaatcggtataaac... tattaaatcttgcaatgtcr 7582
Scoring table: IDENTITY_NUC
 Gappen 10.0 , Gapext 1.0

	28	947.8	12.5	1167	4	US-09-374-766-61
Run on:	29	947.8	12.5	1167	4	US-09-879-844-B-57
Scoring table:	30	932.4	12.3	1158	1	US-08-471-764-1
Gapext 1.0	31	932.4	12.3	1158	2	US-08-471-969-1
Gappen 10.0	32	932.4	12.3	1158	2	US-08-384-177-1
Gapext 1.0	33	932.4	12.3	1158	2	US-08-470-056A-1
Gappen 10.0	34	932.4	12.3	1158	3	US-08-691-593C-1
Gapext 1.0	35	932.4	12.3	1158	3	US-09-200-910-1
Gappen 10.0	36	932.4	12.3	1158	3	US-09-133-411-1
Gapext 1.0	37	932.4	12.3	1158	4	US-09-374-766-1
Gappen 10.0	38	932.4	12.3	1158	4	US-08-979-847B-1
Gapext 1.0	39	914	12.1	9580	4	US-09-949-016-17411
Gappen 10.0	40	895	11.8	68778	4	US-09-949-016-1606
Gapext 1.0	41	863.4	11.4	149971	4	US-09-949-016-1590
Gappen 10.0	42	784	10.3	42917	4	US-09-949-016-13795
Gapext 1.0	43	722.6	9.5	1329	4	US-08-979-847B-108
Gappen 10.0	44	644.4	8.5	2448	3	US-08-691-633C-53
Gapext 1.0	45	644.4	8.5	2448	4	US-09-374-766-53

Total number of hits satisfying chosen parameter: 1202/84 seqs, 818130359 residues

ה'ז

```

1: //cgn2_6_ptodata//1/ina/5A__COMB.seq: *
2: //cgn2_6_ptodata/1/ina/5B__COMB.seq: *
3: //cgn2_6_ptodata/1/ina/6A__COMB.seq: *
4: //cgn2_6_ptodata/1/ina/6B__COMB.seq: *
5: //cgn2_6_ptodata/1/ina/PCTUS.COMB.seq: *
6: //cgn2_6_ptodata/1/ina/backfile1.seq: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

and is derived by analysis of the total score distribution.

סימן

No.	Score	Match	Length	DB	ID	Description
-----	-------	-------	--------	----	----	-------------

2	6124-6	80.8	4	00-07-53-010-A-21	Sequence 1, Append
3	2994-2	39.5	4	US-09-949-016-7417	Sequence 17417,
4	2994-2	39.5	4	US-09-949-016-12249	Sequence 12249,
5	2888-6	38.1	4	US-09-175-928-3	Sequence 3, Append
6	2615-4	34.5	4	US-09-949-016-13002	Sequence 13002,
7	2220-4	29.3	4	US-09-949-016-15858	Sequence 15858,
8	2204-4	29.3	4	US-09-949-016-5640	Sequence 5640,
9	2158-8	28.6	4	US-09-553-080A-26	Sequence 26, Append
10	1837-4	24.2	4	US-08-979-847B-87	Sequence 87, Append
11	1837-4	19.9	4	US-08-949-016-13840	Sequence 1840,
12	1452-4	19.2	4	US-08-979-847B-88	Sequence 88, Append
13	1415-8	18.7	3	US-08-979-563C-57	Sequence 57, Append
14	1415-8	18.7	3	US-03-374-766-57	Sequence 57,
15	1415-8	18.7	4	US-08-979-847B-53	Sequence 53, Append
16	1341-8	17.7	4	US-03-573-080A-54	Sequence 54, Append
17	1341-8	17.5	4	US-08-979-847B-105	Sequence 105,
18	1312-4	17.3	4	US-08-979-847B-207	Sequence 207,
19	1301-8	17.2	4	US-08-979-847B-206	Sequence 206,
20	1200-2	16.9	4	US-08-979-847B-205	Sequence 205,
21	1200-2	15.8	3	US-08-120-653D-1	Sequence 1, Append
22	1087-2	14.3	4	US-08-979-847B-117	Sequence 117, Append
23	1069-4	14.1	4	US-08-691-563C-89	Sequence 89, Append
24	1069-4	14.1	4	US-03-374-766-89	Sequence 89, Append
25	1069-4	14.1	4	US-08-979-847B-83	Sequence 83, Append
26	13.1	2330	4	US-08-120-653D-4	Sequence 4, Append
27	947-8	12.7	3	US-08-691-563C-61	Sequence 61, Append

Query	Match	Best Local Similarity	Score	4648
		95.7%	Pred.	No.
Sequence 21, Appl	Sequence 17417, A			
Sequence 12249, A	Sequence 3, Appl	Matches 6681, Conservative	195,	Mismatches
Sequence 13002, A	Sequence 15888, A			
Sequence 5640, AP	Sequence 26, Appl			
Sequence 87, Appl	Sequence 13840, A	QY	606	GATGGGAAACGTTCCCGCAAGCACAAAC
Sequence 88, Appl	Sequence 57, Appl	Db	4610	GATGGGAAACGTTCCCGCAAGCACAAAC
Sequence 57, Appl	Sequence 53, Appl	QY	666	MCATTGACCTCAGACACTAGAAGAA
Sequence 54, Appl	Sequence 105, App	Db	4670	: : : : :
Sequence 207, App	Sequence 206, App	QY	726	GCACCTCTGACGGAAGATAATAATTACAC
Sequence 205, App	Sequence 1, Appl	Db	4730	GCACCTCTGAGGGAGATAATAATTACAC
Sequence 117, App	Sequence 89, Appl	QY	786	AAGGCAATGAGTGAAGTCATAGTAC
Sequence 89, Appl	Sequence 83, Appl	Db	4790	AAGGCAATGAGTGAAGTCATAGTAC
Sequence 4, Appl	Sequence 61, Appl	Ov	906	CCAGATCCCCGACTCTTCCCAATTAAT

QY	3125	TAATCTTAACTGGRCAATRATTCCAAAGGCCACGGCCCTCAGTGAGGAAYRATCCA	3184	QY	4205	TAYTGGCGTCAGAAACCTRAAGGAGGTGCTTACATCGGGTCATCNAAA	4264
Db	7101	----TACTTAGGGCTAAATTCTCAAGGCCACGGCCCTCAGGACACATCCA	7156	Db	8160	TTATGGCTGAGAAACCTAAAGAGGTGGCAGCTTACATCGGGGTCTCATCGAAA	8219
Db	3185	GCCTTAACTGCTTAACTCTCATCYCAAACCTAAGCACTAAGRRRTCTTGGCT	3244	QY	4265	GGAAGRAMAGGAATASARGAYTCGAAAGAKATATGAGGMAAAAGACTGCA	4324
QY	7157	GCCTTAACTGCTTAACTCTCATCCAAACCTAAAGCACTAAGGGATTCCTGGCT	7216	Db	8220	GAAGAGAAAGGAATAGAGAGAACTGCGAACAGATATGAGGCAAAAGACTGCA	8279
Db	3245	AAYAGGTYTCGCCAAWATGATTCGGTWTGGCRAATACGCCAGGYCATTAWATA	3304	QY	4325	AGGAGGAGCCTCATTAGAATGCTTAAACTCCCTAGTAGGGTAACTCCCTC	4384
QY	7217	ARTAGGTTTCGCCCCAATGATTCCCGGTATGGAACTTCAAGCTTAACTTAATTA	7275	Db	8280	AGCAGGAGCCTCATTAGAATGAGGAACTGCGAACAGATATGAGGCAAAAGACTGCA	8339
QY	3305	CASTPATTAGGAACCTAGAAGCAATTACCCATTARTAGATGAYANCTGAGYR	3364	QY	4385	CGGAAACCAAGGCCACACTAGCAGGAGAACCTCACAGGG-CAG	4443
Db	7276	CACTPATTAGGAACCTAGAAGCAATTACCCATTAGTAGATGAGCTGAGTAG	7335	Db	8340	GGGAAACCAAGGCCACACTAGCAGGAGAACCTCACAGGG-CAG	8399
QY	3365	AATGGCTTCCAGGCCCTTAAGAAGGCCATTAAACCCAGYCCAGTTAGYTTSC	3424	QY	4444	TTTCTCCCTCGGAGCGTTCCCACTGAGAAGGAAATACTTTCCTGACTAT	4503
QY	7336	AAGTGGCTTCCAGGCC-----TAACCCAGCCCAGCTTAACTTAATGTTSC	7381	Db	8400	TTTCTCCCTCGGAGCGCTGAGAACCTTAACACCCCTAGTAGATAGGTAACTCCCT	8459
QY	3425	AACRGGGCAAGACTTTSTYATYRTACAGAAAAAACAGRAAYAGCTCROGAGTC	3484	QY	4504	CATGGAAATFACTTAACCCCTCATCAACCTTCACTTGAGCATGAGCACCA	4563
Db	7382	AACAGGGCAAGACTTTCTCATG-AAAACAGGAAATGCTTCAAGGAGCTTCA	7440	Db	8460	CTATGAAATFACTTAACCCCTCATCAACCTTCACTTGAGCATGAGCACCA	8519
QY	3485	TTACACAGRTCCRAGGGAYGAGCTGCAACCYRTGCRVACTGASTAGGAAYTG	3544	QY	4564	TCATAGGCCAAATCATATTACTGACCAGGCGCTTCATAACTATCAAGCARATAKT	4623
Db	7441	TTACACAGCTTCAAGGAGCACTGCTGCTGCTGACTGTTACTGAACTATG	7500	Db	8520	TTGATGGCCAAATCATATTACTGACCAGGCGCTTCATAACTATCAAGCARAGTAGT	8579
QY	3545	TAATGCGAAAGGGTTGRCYTCATGTTAYGGGTAGTGTGTTAGCAGTYKTAGT	3604	QY	4624	CGGGCTGTGAAKTTGCGCAGARAAATAACCCCTGCTVTGCGAAGCTCTCAGG	4683
Db	7501	TAGTGGCAAAAGGTTGACCTCATTTAGGGTAGTGTGGAGTAGCAGTYKTAGT	7560	QY	4684	AAACAAARAAAGGGCATCCGTGAAARAACTGGCAACTGATGAGGAAAGCCC	4743
QY	3605	CTGAAAGCACTTAAATAACAGGRAGATCTACTGTGAGCACTCTCATGAGTGA	3664	Db	8580	CGGGCTGTGAGTGTGCGCAGAGAAATAACCCCTGCTVTGCGAAGCTCTCAGG	8639
Db	7561	CTGAGAGGTAAATAACAGGRAGATCTACTGTGAGGAGAGATCTCATGAGTGA	7620	QY	4744	AAACCTCAGGGATTCTGATCTACTGTGAGTGTGAGTATACCTTCACGGGTTGGCARAG	4803
QY	3665	AVRGACTACTCTGCTAAGGAGACTGTGGCTGAGACACYGTACTTAATRC	3724	Db	8700	AAACCTCAGGGATTCTGATCTACTGTGAGTGTGAGTATACCTTCACGGGTTGGAG	8759
Db	7621	ATGGCATACTACTGCTAAAGGAGCACTGCTGCTGACTGTTACTTAATGTC	7680	QY	4804	GCCTCCCTGTGAGGAGCAGAAAGGCGCCAAGGGTAATAGGCACTAGTGTGAAATA	4863
QY	3725	AGGCTTAACTCTGAAARGGCCCTGCTGCACTGTGCACTCTAACCCAG	3784	Db	8760	GCCTCCCTGTGAGGAGCAGAAAGGCGCCAAGGGTAATAGGCACTAGTGTGAAATA	8819
Db	7681	AGGCTTAACTCTGAAARGGCCCTGCTGCACTGTGCACTCTAACCCAG	7740	QY	4864	ATTCGGAGATGGACTTCCGGAGCTTACAGGTGACAATAGCCCTGCTTCCAGGCC	4923
QY	3785	YNCATATTCTCCAGACAGTAAAGATAAACTGTCAACAAATTC	3844	Db	8820	ATTCGGAGATGGACTTCCGGAGCTTACAGGTGACAATAGCCCTGCTTCCAGGCC	8879
Db	7741	CCACATTCCTCAGACAGTAAAGATAAACTGTCAACAAATTC	7800	QY	4924	ACAGTACCCAGGGACTATCCAGGGTTAGGATAGATCACTACTGTGCGCTGA	4983
QY	3845	AAACCTATGCCACTCGAGGGACCTTGTAGAGTTCTTGTGAGTCGACTCAAC	3904	Db	8880	ACAGTACCCAGGGACTATCCAGGGTTAGGATAGATCACTACTGTGCGCTGA	8939
Db	7801	AAACCTATGCCACTCGAGGGACCTTGTAGAGTTCTTGTGACTGTCGACCTCAAC	7859	QY	4984	AGGCCACAGTCAGGGAGGTGCGAAATGATGAAATACTGAAAGGACATTTAA	5043
QY	3905	TGTATTAATCTGAACTTCTTGTGAAAAAGGACTCTGAAAGGGGGTATGAGTG	3964	Db	8940	AGCCACAGTCAGGGAGGTGCGAAATGATGAAATACTGAAAGGACATTTAA	8999
Db	7860	TGTATTAATCTGAACTTCTTGTGAAAAAGGACTCTGAAAGGGGGTATGAGTG	7919	QY	5044	AGCAACCCAGGAACCCACCTCACATGCCCTGCTGTGCTTAGCTTAAMAGA	5103
QY	3965	GTCAGGCTTAACTGAAATYTGAAAGTAACTCCAGGAGCTAGTGCTYNGCTR	4024	Db	9000	AGCAACCCAGGAACCCACCTCACATGCCCTGCTGTGCTTAGCTTAAMAGA	9059
Db	7920	GTCAAGTGTAAATGGAACTGCAAGAACTCCACTCTCCAGGAGCTAGTGTGCA	7979	QY	5104	ATCTGCACTTCCCAAAGCAGGACTTAGCCCATACAGAATCTGTGAGGGCC	5163
QY	4025	GCAGAGCTAATGGCTCAVKGKGGACTAGATGGAGAAGGAAATATA	4084	Db	9060	ATCTGCACTTCCCAAAGCAGGACTTAGCCCATACAGAATCTGTGAGGGCC	9119
Db	7980	GCAGAGCTAATGGCTCAVKGKGGACTAGATGGAGAAGGAAATATA	8039	QY	5164	TTATAACCATGACCTGCTGTGAGCCAGACGAGCAACTAGTGTGAGCATCCT	5223
QY	4085	TATAGGACTCTTAACTGAACTGCAAGGAACTGCAATGAGGAACTGCA	4144	Db	9120	TCATACACCATGACCTGCTGTGAGCCAGACGAGCAACTAGTGTGAGAGG	9239
QY	8040	TATAGGACTCTTAACTGCAATGAGGAACTGCAATGAGGAACTGCA	8099	QY	5224	CCTAGCCAAATATCAGCTTAAACATACAGGAACTATCCCTGGAGAGG	5283
QY	4145	AGGGAGTCTTAACTGCAATGAGGAACTGCAATGAGGAACTGCA	4204	Db	9180	CCTAGCCAAATATCAGCTTAAACATACAGGAACTATCCCTGGAGAGG	9239
QY	8100	AGGGAGTCTTAACTGCAATGAGGAACTGCAATGAGGAACTGCA	8159				

Db	9240	GAAAGAACATTCCACCCCTTGTGACATGGTATTAGTCAGTCCTCCCTAATCCC 9299	Db	10319	GTCTGCCTACCCCTAGGAATAATTTCGTCTGTGGTACTCAGCCATCGTTGTAAT 10378
Qy	5344	CATCCCTAGATACTCTGGGAGACCTTACCGCATTTATVACCCAACTCGGG 5403	Qy	6424	GGCTCTCAGAATCTAATGCTCTCTCATCTCTAGTGCCCCYATGRCAATCTACT 6483
Db	9300	CATCCCTAGATACTCTGGGAGACCTTACCGCATTTATVACCCAACTCGGG 9359	Db	10379	GGCTCTCAGAATCTAATGCTCTCTCATCTCTAGTGCCCCYATGRCAATCTACT 10438
Qy	5404	TAAAGTGGCTGGAGTGTCTGGATACTACACITGAGTOAAATCTGGATACTGC 5463	Qy	6484	GRACAGATTATAGTGTATGCTATCTAGGCCGCAAACAAGAGTACCATCTT 6543
Db	9360	TAAAGTGGCTGGAGTGTCTGGATACTACACITGAGTOAAATCTGGATACTGC 9419	Db	10439	GRACAGATTATAGTGTATGCTATCTAGGCCGCAAACAAGAGTACCATCTT 10498
Qy	5464	CAAGGAACCTGAAATTCCAGGAGAACGCTAGTATTCCTGTGAACTCTAGGAT 5523	Qy	6544	CCTTGTATAGGAGCAGAGTGTCTAGGTACTGGCAATGGGACATGGACGGTC 6603
Db	9420	CAAGGAACCTGAAATTCCAGGAGAACGCTAGTATTCCTGTGAACTCTAGGAT 9479	Db	10499	CCTTGTATAGGAGCAGAGTGTCTAGGTACTGGCAATGGGACATGGACGGTC 10556
Qy	5524	TGGCTCTCTCTCAACACACACAGGAAAGTAACTAATCTAACCTCCCATG 5583	Qy	6604	ACCTCACTGCTCTACTACAACTCTCAGAATAGTAACTAATCTAACCTCCCATG 6663
Db	9480	TGGCCCTCTCTCAACACACACAGGAAAGTAACTAATCTAACCTCCCATG 9538	Db	10559	ACCTCACTGCTCTACTACAACTCTCAGAATAGTAACTAATCTAACCTCCCATG 10618
Qy	5584	GCCTCCCTTATCATTTCTCTKAISGTSITYAACCTSTTCACTCACTGCA 5643	Qy	6664	GCGGAGTCCMGTGCTCCCTGCAAGTCACTAACCCCTAGCAGCAGTGCT 6723
Db	9539	GCCTCCCTTATCATTTCTCTKAISGTSITYAACCTSTTCACTCACTGCA 9598	Db	10619	GCGGAGTCCMGTGCTCCCTGCAAGTCACTAACCCCTAGCAGCAGTGCT 10678
Qy	5644	CCCCCTCATGGCGCTGTATGACCAGTAGCTCCCTTACCAAGGTCTATGGAGAT 5703	Qy	6724	AATGCACTGCTCTACTACAACTCTCAGAATAGTAACTAATCTAACCTCCCATG 6783
Db	9599	CCCCCTCATGGCGCTGTATGACCAGTAGCTCCCTTACCAAGGTCTATGGAGAT 9658	Db	10679	AATGCACTGCTCTACTACAACTCTCAGAATAGTAACTAATCTAACCTCCCATG 10738
Qy	5704	CAGGTCCGAAATATGATGCCCATGCTATAGGACTCTTSTAAGGAAACCCAC 5763	Qy	6784	GAAAATGCTGTATATGTTATGCTATCCGGAATGTCAGTCAAGAATGAGT 6843
Db	9659	CAAGGTCCGAAATATGATGCCCATGCTATAGGACTCTTSTAAGGAAACCCAC 9718	Db	10739	GAGAGATGCTGTATATGTTATGTTATGTTATGTTAGGG 10798
Qy	5764	TTCAGTGCACACCCATGCCCCGCACTGTATCACCTGCACTTTCATGAT 5823	Qy	6844	CSAGATGAACTACAGTACAGCAGGAGCTGCT 6903
Db	9779	GCATAATCATTGGCAGGAAATGATTATCTTCTAGTGTCTGGAGACTGGA 9838	Db	10859	CGAGATGAACTACAGTACAGCAGGAGCTGCT 10858
Qy	5884	GGACTCTCTGTTGACTTACTTCACCCAACTGGTAGTCGTCTGATGAGTCA 5943	Qy	6964	CCTCTCTTGGACCCCTATCTTCCTTCTAGGACCTCTGAGCTCTGAGCT 7023
Db	9839	GGACTCTCTGTTGACTTACTTCACCCAACTGGTAGTCGTCTGATGAGTCA 9898	Db	10919	CGCTCTTGTGACCTCTATCTTAACTCTCTGTTACTCTGTTACTCTGTTCTCAGATGCA 10978
Qy	5944	GATCAGGAGAGAAAGATGAAAGAGTAATCTCCAACTCAGGGACATGGC 6003	Qy	7024	GTGTGAACTACAGTGGGCAAGTGGCTGAGCTTCTACCGAGGCC 7083
Db	9899	GATCAGGAGAGAAAGATGAAAGAGTAATCTCCAACTCAGGGACATGGC 9958	Db	10979	GCTGTAACTACAACTGGGCAAGTGGCTGAGCTTCTACCGAGGCC 11038
Qy	6004	ACCTCTGGCCCTCAAAGACTAGATCTCAAACTCATGAAACCTCCGTACCAT 6063	Qy	7084	CTGGAGCGCCGCTGCTGCCAGATCTGATGTTATGACATCAAGGACCCCTCTGAG 7143
Db	9959	ACCTCTGGCCCTCAAAGACTAGATCTCAAACTCATGAAACCTCCGTACCAT 10018	Db	10939	CTGGAGCGCCGCTGCTGCCAGATCTGATGTTATGACATCAAGGACCCCTCTGAG 11098
Qy	6064	ACTGGCTCTGTAACCTTAACTACCCCTACTGSCTCCATGAGSGCTGTGGGCCAA 6123	Qy	7144	GAATCTGAGCTCAACCTCTACTGCGCCAAATCAGGAGCAGTGAAGCT 7203
Db	10019	ACTGGCTCTGTAACCTTAACTACCCCTACTGSCTCCATGAGSGCTGTGGGCCAA 10078	Db	11099	GAATCTGAGCTCAACCTCTACTGCGCCAAATCAGGAGCAGTGAAGCT 11158
Qy	6124	AACCTACTAATGTTGGATGCTCCCTGACTCTGGCATATTTCATCCT 6183	Qy	7204	SGTCGCCAACCTCCACAGGACTTAGTTCTCTGGAGTGGGACTGAGAC 7263
Db	10079	AACCTACTAATGTTGGATGCTCCCTGACTCTGGCATATTTCATCCT 10138	Db	11159	CGTCGCCAACCTCCACAGGACTTAGTTCTCTGGAGTGGGACTGAGAC 11218
Qy	6184	GTACCTGAACTAGGAACTTCAGCAGAAATACGACCTCTGGATTTAGGAA 6243	Qy	7264	AGGACTACTGTTCTCTAGGTAGTAACTGCTTCTGGAGTGGGACTGAGAC 7323
Db	10139	GTACCTGAACTAGGAACTTCAGCAGAAATACGACCTCTGGATTTAGGAA 10198	Db	11219	AGGACTACTGTTCTCTAGGTAGTAACTGCTTCTGGAGTGGGACTGAGAC 11278
Qy	6244	CCTCTGTTCCATCTGGAAATACCCATTACCTAACCTGAGCTTAAATTAGC 6303	Qy	7324	ACATCACCTTAACTGGGAACTGCTTCTGGAGTGGGACTGAGAC 7383
Db	10199	CCTCTGTTCCATCTGGAAATACCCATTACCTAACCTGAGCTTAAATTAGC 10258	Db	11279	ACATCACCTTAACTGGGAACTGCTTCTGGAGTGGGACTGAGAC 11338
Qy	6304	AATRACTACTACACACCACTCCCATGAGTGGAACTCTCCACAAATA 6363	Qy	7384	ACTAAATGCTTAACTGGGAAAGAGGAGTAACTGCTTCTGGAGTGGGACTGAGAC 7443
Db	10259	AATRACTACTACACACCACTCCCATGAGTGGAACTCTCCACACAAATA 10318	Db	11339	ACTAAATGCTTAACTGGGAAAGAGGAGTAACTGCTTCTGGAGTGGGACTGAGAC 11398
Qy	6364	GTCCTGCCTAACCTCTGGAAATATTTCGTGTGTGTGAT 6423	Qy	7444	AGGAGCAGGAGGAACTGCTTCTGGAGTAACTGCTTCTGGAGTGGGACTGAGAC 7503
Db	11399	AGGAGCAGGAGGAACTGCTTCTGGAGTAACTGCTTCTGGAGTGGGACTGAGAC 11458			

QY 7504 CAACCCCTTGGGRCGCCCTCCTTGATAGGAGCTCTTTCATGCATTCACTCT 7563
Db 11459 CAAACCCCTTGGGCCCCCTTGATAGGAGCTCTTCATGCATTCACTCT 11518
QY 7564 ATTAATCTTGCACRGCR 7582
Db 11519 ATTAAATCTTGCACRGCA 11537

RESULT 2
US-09-573-0801-21
Sequence 21; Application US/09573080A
Patient No. 6828097
GENERAL INFORMATION:
APPLICANT: JOAN, KNOELL
APPLICANT: ROGAN, PETER
TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATION
FILE REFERENCE: 30307
CURRENT APPLICATION NUMBER: US/09/573, 080A
CURRENT FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS.: 479
SOFTWARE: Patentin version 3.0
SEQ ID NO. 21
LENGTH: 8523
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: repeat_region
LOCATION: (1)..(8523)
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Jurka, J.; Walichiewicz, J.; Milosavljevic, A.
TITLE: Prototypic sequences for human repetitive DNA
JOURNAL: Journal of Molecular Evolution
VOLUME: 35
ISSUE: 4
PAGES: 286-291
DATE: 1992-10-
DATABASE ACCESSION NUMBER: Database of repetitive elements (rebase)
DATABASE ENTRY DATE: 1996-01-26
US-09-573-0801-21

Query Match 80.8%; Score 6124.6; DB 4; Length 8523;
Best Local Similarity 95.7%; Pred. No. 0; Mismatches 50; Indels 53; Gaps 27;
Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGAAAGGTCCCGGAGAACAAAGGCCCTAAGCGGTTCTGGAAAATGGGA 665
Db 1923 GATGGGAACGTCCTCCCGAACAGAACAAAGGCCCTAAGCGGTTCTGGAAAATGGGA 1982
QY 666 MCACATTGACCTCAGACACTAAGAACAGGACTTATTCCTCTGGATGCGCTG 725
Db 1983 CCACATTGACCTCAGACACTAAGAACAGGACTTATTCCTCTGGATGCGCTG 2042
QY 726 GCACCTCTGAGGAGATTAATTAACACCTCTTCAAGCTAGCYCTTTGAGRA 785
Db 2043 GCACCTCTGAGGAGATTAATTAACACCTCTTCAAGCTAGACCTTTGAGRA 2102
QY 786 AAGGCAAATGGAGTCAGTCCTACAGAACCTTCCTCTTCAAGAACAGACCTAC 845
Db 2103 AAGGCAAATGGAGTCAGTCCTACAGAACCTTCCTCTTCAAGAACAGACCTAC 2162
QY 846 AATTTATGAAAGTGTGAACTGCTTCAAGAACAGCTCACGTCACCTCCATC 905
Db 2163 AATTTATGAAAGTGTGAACTGCTTCAAGAACAGCTCACGTCACCTCCATC 2222
QY 906 CCAGCATCCCCGACTCTTCCCCTTCAACCAAACTGGTCAA 965
Db 2223 CCAGCATCCCCGACTCTTCCCCTTCAACCAAACTGGTCAA 2282

QY 966 AGGAGGATGACAAGGGTAAACAGTGACCAAGAGTGCCAAATTCGCCAAATTGAA 1025
Db 2283 AAGGAGATGACAAGGGTAAACAGTGACCAAGAGTGCCAAATTCGCCAAATTGAA 2342
QY 1026 CCCCTCCMAAGCAGTGGGGAGAGAGAATCGGCCCAAGCAGCTGAGTGATGCTTTT 1085
Db 2343 CCCCT-CCAGCACTGGGGAGAGAGAATCGGCCAGCCAGCTGAGTGATGCTTT 2401
QY 1086 TCTCCCCAGCTTAAGAACATAAAACAGACTTAGGTAATTCTGAGTAAYCTGTG 1145
Db 2402 TCTCCCCAGCTTAAGAACATAAAACAGACTTAGGTAATTCTGAGTAACCTGATGG 2461

QY 1146 CTATATGTTGTTTACAGGGTAGACAATCTTGTATCTGACATGGAGAGATATA 1205
Db 2462 CTATATGTTGTTTACAGGGTAGACAATCTTGTATCTGACATGGAGAGATATA-A 2520
QY 1205 TGTACTGTTAACATCGACACTAACCCAAATGAGAGAATGCGACCATAACTGAGCT 1265
Db 2521 TGTCAGTGTAACTCAGGTTCTCTGTTCTGAGCTCAAGTGTAACTGAGCT 2580
QY 1266 GAGRGTTGGGATCTCTGTTCTGAGCTCAAGTGTAACTGAGCT 1325
Db 2581 GAGGTTTGGCACTCTCTGTTCTGAGCTCAAGTGTAACTGAGCT 2637

QY 1326 AAANGAATGATCCACAGCCAGCARGCAGTCAGTCAGTCTGAGCT 1385
Db 2638 AAAGAGAAATGATGATTCCTCACAGGCCAGAGGCACTTCAGTCAGCTCATGGAGATTA-A 2753
QY 1386 ACAGAAATCAGTACATGGAGATGTTGCTGAGAAATTTGTAATGTTGCTTAA 1445
Db 2695 ACAGAGAAATCAGTACATGGAGATGTTGCTGAGAAATTTGTAATGTTGCTTAA 2753
QY 1446 GGACTAAGAACATASAGAACARTCTYGAATTAATCTCAATGATGTCGAGAAATTGTCACCAAA 1505
Db 2754 GGACTAAGAACATASAGAACARTCTYGAATTAATCTCAATGATGTCGAGAAATTGTCACCAAA 2812
QY 1506 GGGGAAGGAAAGAAATCTACTCGCTTCTGGAGAACTAAGGAGCAATGGAGAC 1565
Db 2813 -GGGAAGGAAAGAAATCTACTCGCTTCTGGAGAACTAAGGAGCAATGGAGAC 2871
QY 1566 GTCGCTCTGACCTGACTCTTCTGGAAACTAACTCTAAGGAAACTAACTCTAAGGAAACTAACTCTAAGGAACTTCTA 1625
Db 2872 GTCGCTCTGACCTGACTCTTCTGGAAACTAACTCTAAGGAACTTCTAAGGAACTTCTA 2931
QY 1626 CTGAGTCAGCTGAGAGCTTAGAACAAACTCTAACAGCTCTGGTAGGCCCGAGCAA 1684
Db 2932 CTGAGTCAGCTGAGAGCTTAGAACAAACTCTAACAGCTCTGGTAGGCCCGAGCAA 2991
QY 1685 ACTTAAAGACCTTATTAATCTGACACCTTCGCTGTTTATATAGAGATCAGGAGAC 1744
Db 2992 ACTTAAAGACCTTATTAATCTGACACCTTCGCTGTTTATATAGAGATCAGGAGAC 3051
QY 1745 AGCGGGAAAGAACAGGAAACGGGTTAAAGAACGGCCCGCTTAGATGACCTCTGG 1804
Db 3052 AGCGGGAAAGAACAGGAAACGGGTTAAAGAACGGCCCGCTTAGATGACCTCTGG 3111
QY 1805 CAAAGTGGACTTGTGAGGCTCTGGAAAGAACGCTCGGCAAAATTGATGCTTAATAGC 1864
Db 3112 CAAAGTGGACTTGTGAGGCTCTGGAAAGAACGCTCGGCAAAATTGATGCTTAATAGC 3171
QY 1865 GCTGTGTCACAGGCGCTACAGGACCTTAAAGAGATGTGTCAGTAGAGTAG 1924
Db 3172 GCTGTGTCACAGGCGCTACAGGACCTTAAAGAGATGTGTCAGTAGAGTAG 3231
QY 1925 CGCCCTTCTGTCACCCCTTATTCAAGGAACACTGAGGCGCACTGCCAG 1984
Db 3232 CGCCCTTCTGTCACCCCTTATTCAAGGAACACTGAGGCGCACTGCCAG 3291
QY 1985 GGAAGAAGTCTTGTGTCAGGCAACTAACAGTGTGTCAGGAGCTGAGGG 2044
Db 3292 GGAAGAAGTCTTGTGTCAGGCAACTAACAGTGTGTCAGGAGCTGAGGG 3351
QY 2045 TGCCTGGGCAAGGCCATCCATGCCATCACCTCACAGGCCCTGGTAGCTTGCAC 2104

Db	3352	TGCTGGGCGAACGCCATCCCA1GCCATCACCTCACAGAGCCCTGGTATCTGACC	3411	Qy
Qy	2105	ATTGAGGCCAGAAGGTGTCRCCSGACACTGGTGGCTCTTAGTCTACTCTC	2164	Db
Db	3412	ATTGAGGCCAGS-AGGTTGTCRCCSGACACTGGTGGCTCTTAGTCTACTCTC	3470	Qy
Qy	2165	TGTCGCCAGAACCTGTCCTCACATCTGACATCTGAGGGGTCTAAGAGGCCA	2224	Db
Db	3471	TGTCGCCAGAACCTGTCCTCACATCTGAGGGGTCTAAGAGGCCA	3528	Qy
Qy	2225	GRCACTAGATCTTYYCCASCCATAGTATGACTGGGAGCTTATCTTCAC	2284	Db
Db	3587	ATGCTTCTAACTGTCGAGATCTGAGGGACACCGTTGTTGNCCTG	3586	Qy
Qy	2345	AAGCAGGGCCATTATACCTGACACATAGGAGGGACACCGTTGTTGNCCTG	2404	Db
Db	3647	AAGCAGGGCCATTATACCTGACACATAGGAGGGACACCGTTGTTGNCCTG	3705	Qy
Qy	2405	CTTGAAGAGGATTAACTCTGACCATAGGAGGGACACCGTTGTTGNCCTG	2464	Db
Db	3706	CTTGAAGAGGATTAACTCTGACCATAGGAGGGACACCGTTGTTGNCCTG	3764	Qy
Qy	2465	GAATGGCCCTCTTCAGTAACTTAAGGATTCCCTTCCCTACAAAGCCA	2524	Db
Db	3765	GAATGGCCCTCTTCAGTAACTTAAGGATTCCCTTCCCTACAAAGCCA	3824	Qy
Qy	2525	GTACCCCCCTOAGACCCCAAGGCCAACAAAGGATTTAGGACTTAAGC	2584	Db
Db	3825	GTACCCCCCTCAGACCCCAAGGCCAACAAAGGATTTAGGACTTAAGC	3884	Qy
Qy	2585	CCAGGCTTAGTAAACCTGCACTTCTCTGAGTAACTCCCTGAGTGGAGG	2644	Db
Db	3938	TAAGGAAACCCACAGACAGTGGAA-GGTTAGTGAAGACTCAGGATACTA	3995	Qy
Qy	2705	CAGAGGCCATTAGTAAACCTGCACTTCTCTGAGTAACTCCCTGAGTGGAG	3937	Db
Db	3996	TGTTGTTCTCTATAGCCAGCTGTACTTACCCCTTATACTCTGCTTCCAAATCCAGA	4055	Qy
Db	4116	TCCGACTCTCAATTCTGCTGACCTTACAGTCCGACCTTCAGATGCTGACA	2824	Qy
Db	4056	GGAGCAGAGTGCTTACASTCCGACCTTACAGTCCGACCTTCAGATGCTGACA	4115	Db
Qy	2825	TCCGACTCTCAATTCTGCTGACCTTACAGTCCGACCTTCAGATGCTGACA	2884	Qy
Db	4176	CTGGTACTTACCCAAAGGGTCAAGGTTAGGGATAGTCCCTATTTGAGAGT	4235	Db
Qy	2945	CCAAGACTTGAGYCRTYMCATACCTGACACTTCTGCTGACCTTAC	3004	Db
Db	4236	CCAAGACTTGAGCCATCTGCATACTGGACA-C-TTGTGCTCTGGTAGTGGATG	4293	Qy
Qy	3065	CCTCGCYACCTGTCGCTACAGGACCCATCTGACCCAGCCTTMAATT	3064	Db
Qy	4294	ACTTTGGCCCTCATCAGAACCTGTCGAGCTTAC	4353	Qy
Db	3125	TAAATCTAGGRCTAATTATCCAAAGCACCARGGCCCTCTGAGGAAVYATCCA	3184	Db
Qy	4414	--TACTTAACTGCTAAATTATCCAAAGGCCCAAGGCCCTCTGAGGAAVYATCCA	4469	Db
Qy	3185	GCTTAACTGCGTTACTCTCTCYAAACCCCTAACACTAGRRRTCTGCGT	3244	Db
Db	4470	GCTTAACTGCGTTACTCTCTCCAAACCTTAAGCACTTAACTGGGATTCCTGGGT	4529	Qy
Qy	3245	AAVAGGTTCTCGGAATGGATTCCAGGTWTGCGRAATAAGCCAGGYATTAWA	3304	Db
Db	4530	AAATGGTTCTCGGAATGGATT-CCCGGTATGGCGAAATAGCCAGGYATTAWA	4588	Qy
Qy	3305	CTSTAATTAAGGAAACTCAGAACCTTACCCATTAACTGAGACATCTCAA	3364	Db
Db	4589	CTSTAATTAAGGAAACTCAGAACCTTACCCATTAGGAGACACTGAGAAGT	4648	Qy
Qy	3365	ADGIGGTTCTAGGCCCTAAAGAAGGCCCTAAACCAAGYCCCTGGTMAHYTCCC	3424	Db
Db	4649	AAAGGGTTCTAGGCCCTAAACCAAGGAACTTAACTGAGATGAGTGAAGMR	4694	Qy
Qy	3425	ACRGSCAAGGCTTSTVATAYTCACAGAAAACAGRAAYGCTTRGGAGTCC	3484	Db
Db	4695	ACAGGGCAAGACTTCTCTGATGAGCAG-AAAACAGGAAATGCTCTGGAGTCC	4753	Qy
Qy	3485	TRACACGRTTCTAGGAGYGCCTCACCTTARTAGATGAGTGAAGTGTG	3544	Db
Db	4754	TRACACGATCGGAGGATGAGCTGCACTCTGGCATACTGAGAATGGATG	4813	Qy
Qy	3545	TGTTGCGAAAGGGTGRYCCTGCTGTTAYGGTAGTGTGTCAGTGGTAKTAT	3604	Db
Db	4814	TAGTGGCAAGGGTGTACTCTGTTACGGAGTGTGAGTGTGAGTGTG	4873	Qy
Qy	3605	CTGAGCGTTAAATATACGGGRAGAGATCTACTGTTGAGACATCTGAKGTG	3664	Db
Db	4874	CTGAGCGTTAAATATACGGGAGATCTACTGTTGAGACATCTGAKGTG	4933	Qy
Qy	3665	ATRGCAACTCTGCTGTAAGGAGACTGTGCTGACTCTGCACTTACCCG	3724	Db
Db	4934	ATRGCAACTCTGCTGTAAGGAGACTGTGCTGACTCTGCACTTACCCG	4993	Qy
Qy	3725	ACGCTTAACTCTGAGGGCAGTGTGCRACTGTGACTCTGCACTTACCCG	3784	Db
Db	5054	CTGAGCGTTAAATATACGGGAGATCTACTGTTGAGACATCTGAKGTG	5113	Qy
Qy	3845	AAACTATGCCCTCGAGGAGCTTGTAGGTTCTGCTGACTCTGCACTTACCCG	3904	Db
Db	5114	AAACTATGCCCTCGAGGAGCTTGTAGGTTCTGCTGACTCTGCACTTACCCG	5172	Qy
Qy	3905	TGTATACCTGAGAAGTCTCTGAGAAGGACTCTGAGAAGGAGCTTGTGAGT	3954	Db
Db	5173	TGTATACCTGAGAAGTCTCTGAGAAGGAGCTTGTGAGAAGGAGCTTGTGAGT	5232	Qy
Qy	3965	GTCTGGTGTAAATGAAATTTGAGAAGTCTCTGAGAAGGAGCTTGTGAGT	4024	Db
Db	5233	GTCTGGTGTAAATGAAATTTGAGAAGTCTCTGAGAAGGAGCTTGTGAGT	5292	Qy
Qy	4025	GAGAACTATGCCCTCTGAGAAGGAGCTTGTGAGAAGGAGCTTGTGAGT	4084	Db
Db	5293	GAGAACTATGCCCTCTGAGAAGGAGCTTGTGAGAAGGAGCTTGTGAGT	5352	Qy
Qy	4145	AGGAACTTCTACTCTGAGAAGGAGCTTGTGAGAAGGAGCTTGTGAGT	4204	Db
Db	5413	AGGGAACTCTCTGAGAAGGAGCTTGTGAGAAGGAGCTTGTGAGT	5458	Qy
Qy	4205	TATGGCGTACGAGAACCTTACAGGAGGTTGAGCTTACCTGCTGAGAAGA	4264	Db
Db	5473	TATGGCGTACGAGAACCTTACAGGAGGTTGAGCTTACCTGCTGAGAAGA	5532	Qy

QY	4265	GGAAGRAAGGSAATASAGRAAYTGCACAGCAKATATGAGCMAAAGAGTCGA	4324	QY	5344	CATCCCTGATACTCTGGAGGCCAACGCTACCGCATTTATACCCCACCTGG	5403
Db	5533	GGAAAGGAAGGGAATAGAGAGACTGCAAGGAGATATTGAGCMAAAGAGTCGA	5592	Db	6513	CATCCCTGATACTCTGGAGGCCAACGCTACCGCATTTATACCCCACCTGG	6672
QY	4325	AGCGAGACCTCATTAGAAACTCTTAACTCTCTTAGTATAGGTATAGCTAACCTTC	4384	QY	5404	TTRAAGTGCTGGAGTGAGTGCTGTTGATACATCACACTGAGTCAAATCTGGAATACCTGC	5463
Db	5593	AGCGAGACCTCATTAGAAACTCTTAACTCTCTTAGTATAGGTATAGCTAACCTTC	5652	Db	6673	TTRAAGTGCTGGAGTGAGTGCTGTTGATACATCACACTGAGTCAAATCTGGAATACCTGC	6732
QY	4385	CGGAAACCAAGGCCAGTAGTCAGCAGGGAGAACAGAATGGGACCTACGAGG-CAG	4443	QY	5464	CAAAGGAACTGAGAACCCAGAGACAGACAGAGGAAACAGAATGGGACCTACGAGG-CAG	5523
Db	5653	CGGAAACCAAGGCCAGTAGTCAGCAGGGAGAACAGAATGGGACCTACGAGG-CAG	5712	Db	6733	CAAAGGAACTGAGAACCCAGAGACAGACAGAGGAAACAGAATGGGACCTACGAGG-CAG	6792
QY	4444	TTTCTCCCTCGGACGCTTAGCACTGAGAAGGAAATACCTTGCCTGCAACTAT	4503	QY	5524	TGGCCCTCTCTCAACACACACAGGAGAGTAATACTAAATCCCCATG	5583
Db	5713	TTTCTCCCTCGGACGCTTAGCACTGAGAAGGAAATACCTTGCCTGCAACTAT	5772	Db	6793	TCGGCTGCTCTCAACACACACAGGAGAGTAATACTAAATCCCCATG	6851
QY	4504	CCCATGGAAATTCTTAAACCTTCATCAAACTTCTTCTTCTGATGAGCACCA	4563	QY	5524	GSCTCTCTTATCATATTTCTKTAAGTGTSTTYACCTSTTACTCTACTGCA	5643
Db	5773	COATGGAAATTCTTAAACCTTCATCAACCTTCTGATGAGCACCA	5832	Db	6852	GCCTCTCTTACATTTCTTACTGTTACTGTTTACCTCTACTGCA	6911
QY	4564	TCAATGGCCAAATCATTTACTGGACCCGCTTCAAACTATCAAGGAAAT	4623	QY	5544	CCCTCTCATGCGCTTATGACCAGTAGCTCCCTTACCAAGAGCTGATGACCCA	5703
Db	5833	TCAGATGCCAAATCATTTACTGGACCCGCTTCAAACTATCAAGGAAAT	5892	Db	6912	CCCTCTCATGCGCTTATGACCAGTAGCTCCCTTACCAAGAGCTGATGACCCA	6971
QY	4624	CAGGGCCTGTTGAAGTGTGCCAARAAATATCCCTGCTTATGCCCAGTCCTCAGG	4683	QY	5704	CAGGGTCCGAAATATGATGCCCCTGCTTATGCCCAGTCCTCAGG	5763
Db	5933	CAGGGCCTGTTGAAGTGTGCCAAGAAATATCCCTGCTTATGCCCAGTCCTCAGG	5952	Db	6972	CAAGCTGCTCTCTCAACACACACAGGAGGAAGTAATAATCTAAT-CCCATG	7031
QY	4684	ARACAAARACAGGCCATTACCTGARAACTGGCAGCTTACCCACAAGCC	4743	QY	5764	TTACTGCCACACCTATGCCCGGACTCTACACTGCTCCCTTCTCAT	5823
Db	5953	AGACAAAGAACAGGCCATTACCTGAGAACAGCTGAGCTGACAGCTGAGG	6012	Db	7032	TTCACTGCCACACCTATGCCCGGACTCTACACTGCTCCCTTCTCAT	7091
QY	4744	AACCTCAGGATTCACTACTACTAGCTGCTTACACTTACCTTACGGTTEGGCARAG	4803	QY	5824	GCATACATCATTTGACAGGAGTAGCTCCCTTACACTCTGCTCCCTTCTCAT	5883
Db	6013	AAACCTCAGGATTCACTACTACTAGCTGCTTACACTTACGGTTEGGCARAG	6072	Db	7092	GCATACATCATTTGACAGGAGTAGCTCCCTTACACTCTGCTCCCTTCTCAT	7151
QY	4804	GCCTTCCCTGTAGGACAGAAAGGCCAGAGGATAAGGGACTACTCTGAAATA	4863	QY	5884	GTCACTCTGTGACTTACCTCACCCATGCTTAACTGGGTGAGTCATGCA	5943
Db	6073	GCTTCCCTGTAGGACAGAAAGGCCAGAGGATAAGGGACTACTCTGAAATA	6132	Db	7152	GTCACTCTGTGACTTACCTCACCCATGCTTAACTGGGTGAGTCATGCA	7211
QY	4864	ATTCAGCATCTGCACTTCCGAGGCTTAAGAGTGTCAATGCCCTTCCAGGCC	4923	QY	5944	GATCAGCGAAGGAAGAACAGTAAGGAAGTAACTCTCCACACTCGSGGTACATGCA	6003
Db	6133	ATTCAGATTGCGACTTCCGAGGCTTAAGAGTGTCAATGCCCTTCCAGGCC	6192	Db	7212	GATCAGCGAAGGAAGAACAGTAAGGAAGTAACTCTCCACACTCGSGGTACATGCA	7271
QY	4924	ACGTAACCAAGGAGTATCCCGCGGTAGGTTACGAGTGTCAATGCCCTGCA	4983	QY	6004	ACCTCTGCCCTACAGAGGACTAGTCTCAACACTACATGAAACCTCCGTCACCAT	6063
Db	6193	ACGTAACCAAGGAGTATCCCGCGGTAGGTTACGAGTGTCAATGCCCTGCA	6252	Db	7272	ACCTCTGCCCTACAGAGGACTAGTCTCAACACTACATGAAACCTCCGTCACCAT	7331
QY	4984	AGGCACAGCTCTCGAGGAGGTGAGAAATGAGTAACTCAAGGACATCTAAA	5043	QY	6064	ACTCGCTGTGACTTACCTACACCCCTACTGGCTCATGAGGTCCTGGCCCA	6123
Db	6253	AGGCACAGCTCTCGAGGAGGTGAGAAATGAGTAACTCAAGGACATCTAAA	6312	Db	7332	ACTCGCTGTGACTTACCTACACCCCTACTGGCTCATGAGGTCCTGGCCCA	7391
QY	5044	AGGAAACCCAGGAACCCCTCACATGGCTTCTGTGCTTATGCTTAAGGA	5103	QY	6124	ACCCCTCTAACCTGCTGAGTAGCTCCCTGAAACTCARECCATGTTCATCCT	6183
Db	6313	AGGAAACCCAGGAACCCCTCACATGGCTTCTGTGCTTATGCTTAAGGA	6372	Db	7392	ACCCCTCTAACCTGCTGAGTAGCTCCCTGAAACTCAGGCTACCTGGCTCATGAGGTCCTGGCCCA	7451
QY	5104	ATCGCAACTTCCCAAAGAGCAGGACTAGCCATAGGAATGCTGTGAGBAGGCC	5163	QY	6184	GTACTCTGAACTGAGGAACTCAGGAGAGAACTCTGCTTGTGAGGAGAGG	6243
Db	6373	ATCGCAACTTCCCAAAGAGCAGGACTAGCCATAGGAATGCTGTGAGBAGGCC	6432	Db	7452	GTACTCTGAACTGAGGAACTCAGGAGAGAACTCTGCTTGTGAGGAGAGG	7511
QY	5224	CCTTAGGCAATCTCCCAAAGAGCAGGACTAGCCATAGGAATGCTGTGAGGAGG	5283	QY	644	CCCTCTGTCTGACTCTGCTCTCCCTGAGCTACCTGGTAAATATTAGC	6303
Db	6493	CCTTAGGCAATCTCCCAAAGAGCAGGACTAGCCATAGGAATGCTGTGAGGAGG	6552	Db	7512	CCCTCTGTCTGACTCTGCTCTCCCTGAGCTACCTGGTAAATATTAGC	7571
QY	5284	GAAAAGAACTATTCCACCCWGGTGCACATGTGTTAGTCAAGTCCTTCTPATTCC	5343	QY	6344	ATACTCTACACACACACCTCCCATGCTGAGGAGAGAACTCTGCTTGTGAGGAGG	6363
Db	6553	GAAAAGAACTATTCCACCCWGGTGCACATGTGTTAGTCAAGTCCTTCTPATTCC	6612	Db	7572	ATACTCTACACACACCTCCCATGCTGAGGAGAGAACTCTGCTTGTGAGGAGG	7631
QY				Db	6364	GTCTGCCTACCTCAGGAAATTTCTGCTGTCAGTCCTCTGAGCTTGTGTTGTGAGGTT	6423
Db				Db	7632	GTCTGCCTACCTCAGGAAATTTCTGCTGAGCTTGTGAGGAGAGAACTCTGCTTGTGAGGAGG	7691
QY				Db	6424	GGCTCTCAGAACTCTGCTCTCATCTCTAGTGCCCTTGTGAGCTTGTGAGGAGG	6483

NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO: 17417
; LENGTH: 77772
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(77772)
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17417

Query Match 3.9-5% Score 2994.2; DB 4; Length 77772;
Best Local Similarity 85.5%; Pred. No. 0; Mismatches 362; Indels 69; Gaps 23;
Matches 3418; Conservative 149; Mismatches 362; Indels 69; Gaps 23;

QY 1812 ACTTGAGGCTCTGAAAGGGAAAGCTGCGCAATTGATGCTAATGGGCTGTCT 1871
Db 19931 ATTTGAGGCTCTGGAAGAAGGGAAAGCTGCGCAATTGATGCTAATGGGCTGTCT 19872
QY 1872 TCCAGTGGCGCTACAAGGACACTTAAGAAAAGATGTCAGTGAAGTAGAAGTGGCGCCC 1931
Db 19871 TCCAGTGGCGCTACAAGGACACTTAAGAAAAGATGTCAGTGAAGTAGAAGTGGCGCCC 19812
QY 1932 TCTGCCATGCCCTTATTCAAGGGATCACTGAAAGCCACTGCCCCAGGGGCAA 1991
Db 19811 -TCATCCATGCCCTCATGTCAGGAACTTAAGAAAT 8111
QY 1992 GGTCTTGTGAGTCAAGGCACTAACAGGATGATCCAGCAGGAGCTGGGTCTGG 2051
Db 1992 GGTCTTGTGAGTCAAGGCACTAACAGGATGATCCAGCAGGAGCTGGGTCTGG 19753
QY 19752 GTCTCTGAGTCAAGGCACTAACAGGATGATCCAGCAGGAGCTGGGTCTGG 19693
Db 2052 GGCAGAGGCCACTCCATGTCAGGAACTGCGCCAGGGGAA 19650
QY 2112 GCGCAGGAAAGT--TGTCTCTGAGCAGCTGTCAGGCGCTCTTCTGTCAGTCTCTGTC 2168
Db 19649 GCGCAGAAGGGT--TGTCTCTGAGCAGCTGTCAGGCGCTCTTCTGTCAGTCTCTGTC 19590
QY 2169 CGGAAACGTCCTCCATGTCAGGAACTGCGCCAGGGGCTTAAGGGCAGTC 2228
Db 19692 GCGCAGGCGCCAT-----GCCCGAGTGTCTGACATGG 19533
QY 2211 GCGCAGGAAAGT--TGTCTCTGAGCAGCTGTCAGGCGCTCTTCTGTCAGTCTCTGTC 216
Db 19753 CTGACAACTCTCTCCAGATCTGACTACC--GASGGGTCTCPAGGATAGGAGCTCA 19475
QY 2229 CTAGATACCTTCTCCAGGACTCTGAGTGTATGACATGGACCCCTCTGAG 2288
Db 19532 CTAGATAC-TTCTCCAGGACTCTGAGTGTG-ACTGGGGACTCATCTTCTGAC 1945
QY 2289 TTCTCAATATGCTGAGGCCACTACCTGTTAGGAGACATCTGAGAACAG 2348
Db 19474 CTTCTCAATATGCTGAGGCCACTACCTGTTAGGAGACATCTGAGAACAG 19415
QY 2349 AGGGGCAATTACCTGAGCATAGGAGGGACACCCCTTGTAGGGAGATCATCTGAGAACAG 2408
Db 19414 AGGGGCAATTACCTGAGCATGGAGGGACACCCCTTGTAGGGAGATCATCTGAGAACAG 19357
QY 2409 AGGAGGAAATTACTCTGAGGCTGGCAACAGGAGCATATGGACGACCAAGAT 2468
Db 19356 AGGAGGAAATTACTCTGAGGCTGGCAACAGGAGCATATGGATGAG-GGAGAT 19298
QY 2469 GCCCCTCCCTGCTGAGTAACTAAGGATCTCCCTTCCCTACCAAGGGATAC 2528
Db 19297 GCGCGCCCTGCTGAGTAACTAAGGATCTCCCTTCCCTACCAAGGGATAC 19238
QY 2529 CCCCTCAGACCCAAAGGCCAACAGGATCCAAAGGATGTTAGGACTAAAGCCAA 2588
Db 19237 CCTCTAGACCCAAAGGCCAACAGGACTCCAAAGATGTTAGGACTAAAGCCAA 19178
QY 2589 GGCTTAGTAAACCATGCTACCTCTGAGTAATCCGTTAGTGGATGGAGGACAA 2648
Db 19177 GGCTTAGTAAACCATGCTACCTCTGAGTAATCCGTTAGTGGATGGAGGACAA 19125
QY 2649 GAACCCAGTGGAGCTGGAGGGTTAGTGGAGAAGTCTCAGGATTAATGGGGCGTT 2708

RESULT 3

US-09-949-016-17417/C

; Sequence 17417; Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIORITY NUMBER: 60/241,755
; PRIORITY FILING DATE: 2000-10-20
; PRIORITY APPLICATION NUMBER: 60/237,768
; PRIORITY FILING DATE: 2000-10-03
; PRIORITY APPLICATION NUMBER: 60/231,498
; PRIORITY FILING DATE: 2000-09-08
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 7797
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1...7797)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-12249

Query Match 39.5%; Score 2994.2; DB 4; Length 7797;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 3418; Conservative 149; Mismatches 362; Indels 69; Gaps 23;

QY 1872 TCCAGTGGGTACAGGAACTTAAAGATTCTCAAGTAGAAGTCAGCGCC 1931
Db 19875 TCGAGTGTTCTACAGAACCTTAAAGATTCTGAATAGATAAGTCGCC 19816
QY 1932 TTGTCTCATGCCCTCATCTCAGGAACTCTGGAAGGCCACTGCCAGGGACA 1991
Db 19815 -TCATCATGCCCTCATCTCAGGAACTCTGGAAGGCCACTGCCAGGGATA 19757
QY 1992 GGTGTTGAGTGTGAAAGGAACTTACCGATGATGCCAGGAGGCTGCG 2051
Db 19756 GGTCTTTCAGTCAAGAACCTACAGATAGCCAGGAGGACTGGGCC 19697
QY 2052 GGCAGAGGCCATCCCATCCATCACCTCACAGAGCCCTGGGATGCTGAC 2111
Db 19696 GGCAGAGGCCAT-----GCCAGGATATGCTGACATGGCG 19654
QY 2112 GCGAGGAGGT--TGTCTCTGACCTGAGCTGGCTGGGCTCTAGCTCTACTCTCTGTC 2168
Db 19653 GCGAGAAGTTAACTGTCTCTGACCTGAGCTGGCATGGCTCTAGCTCTGTC 19594
QY 2169 CGCGACACTGTCTCGATGTCATTTGAGGAGGAGGCTTAAAGGCGAGTC 2228
Db 16205 --ATGGCCCCCGCTCATTTCTCTTACTGTGTCTACCCCTTCACTACA 16148
QY 5639 CTGCAACCCTCATGCCGTGTATGCCGTAGCTCCCCTACCMAGATTCTATGA 5698
Db 16147 CCTCACCCCTCCATGCCGTATACCTACCGTAGCTCCCTTACCAAGACTCTATGA 16088
QY 5699 GAATCAGCCTCCCGGAATATGTGAGCCATCGTATAGAGCTTAAAGGAAACC 5758
Db 16087 GAATCGGCTCCCGAATATGTGAGCCATCGTATAGAGCTTAAAGGAAACC 16028
QY 5759 CCACCTCTCGCCACCCATATGCCGCACTGCG 5796
Db 16027 CCACTTCACATCCACCCATATGCCGCACTGCG 15990

RESULT 4
US-09-949-016-12249/C
; Sequence 12249, Application US/09/949,016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

QY 2409 AGGAGGATTAATCCGAGTCAGGACAGGACATATGGACGCCAAGAAGT 2468
Db 19360 AGGAGGAAATTATCTGAGTCAGGACAGGACATATGGATGAG-GGAGAAT 19302
QY 2469 GCCCGTCTGTCAGGAACTAAGGAGTCCACTCTTCTTACAAAGGAGTAC 2528
Db 19301 GCCGTCCGTCAGTAACTAAGGATTCCGCTCTTCCACCAAGGAGTAC 19242

QY	2529	CCCTCAGACCAAGGCCAACAGGATTCCAAAGATGTTAGGACTTAAGCCCA	2588	QY	3606	TGAAGGNTAAATACTACGGRIGAGATCTACTGTGTCGACITCTCATGKGAA	3665
Db	19241	CCTCTTAAACCCGAGGCCAACAGGACTCCAAAGATGTTAGGACCTAAAGCCCA	19182	Db	18114	TGAGGNTAAATACTACGGRIGAGATCTACTGTGTCGACITCTCATGKGAA	18115
QY	2589	GGCTTAGTAACCATGCTAACTCCCTGCACTAATCGTGAGTGGAGGCACA	2648	Db	18114	YRGCAACTACTGCTAAAGGAGCTTGTCGACACTGCTGCACTTAATRTCA	3725
Db	19181	GGCTACTAAACCATGCTAACTCCCTGCACTAATCGTGAGTGGAGGCACA	19129	Db	18114	CGGCATACTACTGCTAAAGGAGCTTGTCGACACTGCTGCACTTAATRTCA	18055
QY	2649	GAACCGAGTGGACAGTGGAGGTAGTGCAGATCTGGATTATGAGGGCTT	2708	QY	3726	GGCTCTAACTTGCTGAAAGGCAGTGTGCACTGTCGACACTCTAACCGY	3785
Db	19128	GAACCCAAATGACAGTGGAAAGTACTGCAAGATCTGGATTATGAGGGCTT	19071	Db	18054	GGCTCTAACTCGTGAAGGGCAGTGTGCACTGTCGACACTCTAACCGY	17995
QY	2766	GAAGCAGAGTGGTTACCCAGCTGACTCTAGGCCTT-TATAGTGTGTCGAC	2825	QY	3786	CNCATTCCTCCAGAAATGAGATAAATGAGATCTGCACTTCTGTCAC	3845
Db	19010	GAAGCAGAGTGGTTACAGTCTGGACCTTAAGGATGCTT-TATAGTGTGTCAC	18951	Db	17994	CACATTCTTCAGAATGAGAAAGAGATGAAACTGTCGACACTCTAACCGY	17935
QY	2826	CCTGACTCTCAATTCTGTGTCCTTGAGAAGTACTCAACCCARCATCTAAC	2885	QY	3846	AACCTATGCTCTGAGGACTCTGAGAAGAGGCTTGAAGAAGGGTATGCTG	3965
Db	18950	CCTGACTCTCAATTCTGTGTCCTTGAGAAGTACTCTAACCGAC	18891	Db	17934	ACACTAGGCTCTGAGGAGCTCTAGAGTCTCTGACTGACCTGACCT-TCAC	17876
QY	2886	TGGACTTTTACCCCAGGGTCAGGGATAGYCCCATPTTGGCCAGCATAGCC	2945	Db	17906	TGATPCTGATGGAGTCTTGAGAAGAAGGCTTGAAGAAGGGTATGCTG	3905
Db	18890	TGGACTTTTACCCCAGGGTCATOGAACCCCCATTTGGCCAGCATAGCC	18831	QY	17875	TGATPCTGATGGAGTCTTGAGAAGAAGGCTTGAAGAAGGGTATGCTG	3965
QY	2946	CAAGACTGAGCARTYNTCATACCTGAGACTCTGTCCTGTCAGTGGATGATTA	3005	QY	3966	TCAGTGTATGGAATAYTGAAGTAACTCTCCCTACTCCAGGAACTGAGTC	4025
Db	18830	CGAGACTTGTGAGCAGTCTCATACTCTGGACTCTGCTCTGGATGATTA	18771	Db	17815	TCAGTGTATGGAATACTGAGTAACTGAACTTACCTCCCTACTCCAGGAACTGAGTC	17755
QY	3006	CTTTRGCYGYCRTTCAGAAACCTGTGCCATCAGCCACCAAGRCRCTTMAATC	3065	QY	4026	CAGAACTTAATGCCCTACTCAGGACTCTAATAGGAGAAGGAAATAT	4085
Db	18770	CTTTAGGCCACCATTCAGAACCTGTGCCATCAGCCACCAAGTGTCTTAAC	18711	Db	17755	CAGAACTTAATGCCCTACTCAGGACTCTAATAGGAGAAGGAAATAT	17696
QY	3066	CTCGCYAACCTGTGCTCAWGGTTCAACASARAGCTACRTCTGCTACAGAGGT	3125	QY	4086	ATACAGACTTARATGTCYACCTGTGAGGAACTTARAGAGGAGAAGGAAATAT	4145
Db	18710	ATGGCACCTGTGCTACAGTTCAGACCAAGGCTCTAGCTCTACAGAGGT	18651	Db	17695	ACACAGACTTAAAGTGTACTCTACAGGAACTGAGGAAAGGAAATAT	17636
QY	3126	AAACTCTGAGCTAAATATTCAGGAGGACCTGAGGAAAGGAACTGAGGAAAG	3185	QY	4146	GGAAATCTCTACTCYGAGAACCTATCACACATCAGGAGCCATTAGGAA	4205
Db	18650	AAATCTTAGGGCTAAATATTCAGACACACAGGGCCCTCAGTGGAGATCAG	18591	Db	17635	GGGAATCTCTACTCYGAGAACCTATCACACATCAGGAGCCATTAGGAA	17576
QY	3186	CCTRACTGSGTATCTCTCYAAACCTAAAGCACTAAGCACTAAGRRRTCTGCR	3245	QY	4206	TAYTGGCWGTACAGAAACCTTACAGGAGGAGTCTACACTGCGGGTGTAC	4265
Db	18590	CCTRACTGSGTATCTCTCYAAACCTAAAGCACTAAGCACTAAGRRRTCTGCR	18531	Db	17575	ATTCGCTGTCAGAAACCTAAAGGAGTGGCAGCTTACACTGCGGGTGTAC	17516
QY	3246	AYAGGYTCTCCCGAAWATGGATCCCGAAGSTWTGGCRAATAGCAGGTCATWATC	3305	QY	4266	GAAGGRAGGGAAATASAAGRAYTGCAGAAKATATGAGGMMAAAGAGSTGCAA	4325
Db	18530	ACAGCTTCTCCGGAAATGGATTCCAGACCAAGGCTCTAGTGGAGATACAC	18472	Db	17515	GAAGGAGAAGGAAATAGAGGGACCGCGGAGGAACTGAGGAAAGGCAA	17456
QY	3306	ASTAATGAGAACTCAGAGCCATACCCATTARTAGATGAYMAGCTGAAGY	3365	QY	4326	GGCAGGACCTCCATTAGAAATGCTTAACTCTCCCTAGTAGGGTAATCCTCC	4385
Db	18471	ACTAATTTAGGAAACTCAGAGGAACTACCCATTAGTGAAGAGGAACTGAGGAA	18412	Db	17455	GGTGGGACCTCCATTAGAAATGCTTAACTCTCCCTAGTAGGGTAATCCTCC	17396
QY	3366	AGTGGCTTCAGGCCCTAAAGAGGCCATAACCAAGGCCAGTGTAGTGTAG	3425	QY	4386	GGGAAACCCAGGCCAGGACTCTAGCAAGGAGAAGGAAACTCAGAGGAACT	4445
Db	18411	AGTGGCTTCAGGCCCTAAAGAGGCCCTAAAGAGGCCATAACCAAGGCCAGTGTAG	18354	Db	17395	AGGAAACCCAGGCCAGGACTCTAGCAAGGAGAAGGAAACTCAGAGGAACT	17336
QY	3426	ACRGCCACAGCTTISTTATYRTCAAGAAAGAACAGRAAYAGCTCTGGAGTC	3485	QY	4446	TTCCTCCCTGGGAGGTTAGCCACTGAGGAAACTCTAGCTGCTGCACTATC	4505
Db	18353	AAGGGCACACCTTCTTAAACCTTCATCAGACCTTCTAGGATCTAC	18295	Db	17335	TCTTCCTCTAGGAGGCTGAGGAAAGGAAACTCTAGCTGCTGCACTATC	17276
QY	3486	TACAGAGTCRAGGAGGAGCTGAGGCTGAGGAACTCTAGTGGAGTC	3545	QY	4506	ATGGAATCTACTAAACCTCTCATAAACCTTCACTTAGGCATGAGCACC	4565
Db	18294	TACAGAGTCRAGGAGGAGCTGAGGAACTCTAGTGGAGTC	18235	Db	17275	ATTCGAACTACTAAACCTCTCATCAGACCTTCTAGGATCTAC	17216
QY	3546	AGTGCACAGGGTTCYCTATTGTTAYGGTAGGGTAGGAGCAGTGTAGTAC	3615	QY	4566	ARATGGCAACATATTACTGGAGGAAACTCTAGCTGCTGCACTATC	4625
Db	18234	AGTGCACAGGGTTCYCTATTGTTATGGTTATGGTAATGGCAGCAGCTTAGTAC	18175	Db	17215	AGAGGCCAACTCTTAACTCTGTCAGGAGTCAGCAGTC	17156
QY				QY	4626	GGGCTGTGAGTGTGCCRARAATAATGCCCTGCTVATGCGCACTCTGCGG	4685
QY				QY	17155	GGGCTGTGAGTGTGCCRARAATAATGCCCTGCTVATGCGCACTCTGCGG	17196
QY				QY	4686	AACAAACAGGCCATTACCTGCRARAABACTGGCAACT-GATT	4744

Db 16031 CCACCTTCACCATCCACACCCATATGCCCTGCACTTC 15994

Db 17095 AACRAGAACAGGCCATTACCCAGGAGAACTGGCACTAGATTTACCCATGCCA 17036

Oy 4745 AACCTCAGGATTTCAGTTCTATACTGTGGPATARCTTTCACGGGTGGCARGG 4804

Db 17035 AATCTCAGGATTTCAGTACTACTAGTGTGAGATACCTTCACGGGTGGCARGG 16976

Oy 4805 CCTTCCTCTGTAGGACAGAAGAGGCCAAGGGTATAAGGCCTAGTCAGAATA 4864

Db 16975 CCTTCCTCTGTAGGACAGAAGAGGCCAAGGGTATAAGGCCTAGTCAGAATA 16916

Oy 4865 TTCCAGATTCGGACTTCCTCGAGCTACAGTGACATGAGCTACAGTGACATGAGCT 16924

Db 16915 TTCCAGATTCGGACTTCCTCGAGCTACAGTGACATGAGCTACAGTGACATGAGCT 16856

Oy 4925 CAGTAACCCAGGGATATCCAGGG---TTAGTATAGATACTATACACTGCSC 4980

Db 16855 CAGTAACCCAGGGATATCCAGGG---TTAGTATAGATACTATACACTGCSC 16796

Oy 4981 TGAGGCCACAGTCTCAGGAAGTCGGAAATGAGTAATGAGGACATCA 5040

Db 16795 TGGAGCCACAACTCTCAGGAACCCACCTCGCTGGCTGTGTCCTAGCTTACAGGT 16736

Oy 5041 AAAAGCAACACCCAGGAACCCACCTCACAGGCTTAGCCATAGGAATATGAGAACACTCA 5100

Db 16735 AAAAGCAACACCCAGGAACCCACCTCGCTGGCTGTGTCCTAGCTAACACATCA 1676

Oy 5101 AGATCTGCACTTCCCAAAGCAGGCTAGCCATAGGAATATGAGAACACTCA 5160

Db 16675 AGATACAAACACTCCCCAAAAGCAGGACTAACCAATAGGAATATGCTGATGAG 16616

Oy 5161 CCCCTCATACAAATGACCTTGCTTGACCA AGACACCAACTTGTGAGACATC 5219

Db 16615 CCCCTCTTACCAATGACCTTGCTTGAGACATC 16556

Oy 5220 ACCCTCTTACCAATGACCTTGCTTGAGACATC 5279

Db 16555 ACCCTCTTACCAATGACCTTGCTTGAGACATC 16496

Oy 5280 GAGGAAAGAACATTCGCCCGWGTGAGATGGATTACTCAAGTCCCTCYCTCAAT 5339

Db 16495 GAGGAAAGAACATTCGCCCGWGTGAGATGGATTACTCAAGTCCCTCYCTCAAT 16439

Oy 5340 TCCCATCCCTAGATACATCTGGAGGCCCTACCCATTTATTAACCAACT 5399

Db 16438 TCCCACCCCTAGATACATCTGGAGGCCCTACCCATTTATTAACCTGAGT 16379

Oy 5400 GCGGTTAACCTGGCTGGAGAGCTGGATACTACATCTGGATACTTCGGATA 5459

Db 16378 GCAGTTAACCTGGCTGGAGAGCTGGATACTACATCTGGATACTTCGGATA 16330

Oy 5460 CTGCCAACGAACTCTGGCTGGAGAGCTGGATACTACATCTGGATACTTCGGATA 5519

Db 16329 CGCCCAAGAACCCAAATTCTAGGAGAACCTGGATACTTCGGATACTTCGGATA 16270

Oy 5520 GATTGGCTCTGCTCTCACACACACCCAGGAGAACTAAATATAATCCCA 5579

Db 16269 GATCTGACACTGTCTCTCACAGGAGAACCTGGAGAAAGTAATGAGCTC 16210

Oy 5580 CATGGSCTCTTATCATATTTCCTKTAAGTGTSTTY-A-CCTSTTCATCTCA 5638

Db 16209 -ATGGCCCCCCCCGTCATATTTCCTTACTGTGTCTACCCCTTCATCTCA 16152

Oy 5639 CTGACACCCCTCCATGCCGCTGATGAGCACTAGTCCTCTYACAGAGCTTCATGGA 5698

Db 16151 CCTCAACCCCTCCATGCCGCTGATGAGCACTAGTCCTCTYACAGAGCTTCATGGA 16092

Oy 5699 GAATGAGGCTCCGGAAATATGATGCCCATGCTGAGCTCTTAAAGGAGACC 5758

Db 16091 GAATGAGGCTCCGGAAATATGATGCCCATGCTGAGCTCTTAAAGGAGACC 16032

Oy 5759 COACCTTCATGCCAACCCATATGCCCTGCAACTCC 5796

Db ; Sequence 3, Application US/0175928A

Oy ; GENERAL INFORMATION:

Db ; APPLICANT: Jacobs, Kenneth

Oy ; APPLICANT: McCoy, John M.

Db ; APPLICANT: Lavallie, Edward R.

Oy ; APPLICANT: Collins Racine, Lisa A.

Db ; APPLICANT: Evans, Cheryl

Oy ; APPLICANT: Merberg, David

Db ; APPLICANT: Treacy, Maurice

Oy ; APPLICANT: Genetics Institute, Inc.

Db ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

Oy ; FILE REFERENCE: 6006B.AU172A

Db ; CURRENT FILING DATE: 1998-10-20

Oy ; NUMBER OF SEQ ID NOS: 62

Oy ; SOFTWARE: Patentin Ver. 2.0

Oy ; SEQ ID NO: 3

Oy ; LENGTH: 2946

Oy ; TYPE: DNA

Oy ; ORGANISM: Homo sapiens

Oy ; US-09-175-928-3

Query Match 38.1% Score 2888.6; DB 3; Length 2946;

Best Local Similarity 98.2%; Pred. No. 0; Matches 2872; Conservative 49; Mismatches 3; Indels 1; Gaps 1;

Oy Db 4658 CTGCTTATGCCAGCTCTCTGAGGAGAACCTGGCCATTACCTGGAGAGAC 4717

Oy Db 6 CTGCTTATGCCAGCTCTCTGAGGAGAACCTGGCCATTACCTGGAGAGAC 65

Oy Db 4718 TGGCAATCTGATTACCCAAAGGCCAACCTGGAGAACAGAGAACAGGCCATTACCTGGAGAGAC 4777

Oy Db 66 TGGCAATCTGATTACCCAAAGGCCAACCTGGAGAACAGAGAACAGGCCATTACCTGGAGAGAC 125

Oy Db 4778 TARACTTTCAGGGTGGCARAGGCCTCCCTGGAGAACAGAGAACAGGCCATTACCTGGAGAGAC 4837

Oy Db 126 TAGATCTTCACGGTTGGAGAACAGGCCTCCCTGGAGAACAGAGAACAGGCCATTACCTGGAGAGAC 185

Oy Db 4838 TAATAGGACTAGTCTGAAATAATCCAGATGGTGGACTCTCCCGAGCTTACAGA 4897

Oy Db 186 TAATAGGACTAGTCTGAAATAATCCAGATGGTGGACTCTCCCGAGCTTACAGA 245

Oy Db 4898 GTGACATAGGCCCTGCTTCCAGGCCAACCTGGAGGATTCCTGGAGGTTAGGTA 4957

Oy Db 246 GTGACATAGGCCCTGCTTCCAGGCCAACCTGGAGGATTCCTGGAGGTTAGGTA 305

Oy Db 4958 TAGATATCCTTACCTGGCCCTGAGGCCAACCTGGAGGATTCCTGGAGGTTAGGTA 5017

Oy Db 306 TAGATATCCTTACCTGGCCCTGAGGCCAACCTGGAGGATTCCTGGAGGTTAGGTA 365

Oy Db 5018 ATGAAATCTAACAGGACACTAAAGAACCCAGGAAACCCACTCACGGCTG 5077

Oy Db 366 ATGAAATCTAACAGGACAATCTAAAGAACCCAGGAAACCCACTCACGGCTG 425

Oy Db 5078 YTCTGTGCTTATGCTTAAAGATCTGCACTTCCTCCAAAGAGGAGCTTACGGCT 5137

Oy Db 426 CTCCTGCTTATGCTTAAAGATCTGCACTTCCTCCAAAGAGGAGCTTACGGCT 485

Oy Db 5138 CATGGAAATGCTGTGAGGCTCTCATACCAATGACCTTGCTGACCAAGAC 5197

Oy Db 486 CTCCTGCTTATGCTTAAAGATCTGCACTTCCTCCAAAGAGGAGCTTACGGCT 545

Oy Db 5198 AGCCACTTAGTGTGAGCACTCACCTCTAGGCAATATGACCAAGGTCTTAAACATT 5257

Oy Db 546 AGCCAATTTGTCAGACATCACCTCTAGGCAATATGACCAAGGTCTTAAACATT 605

RESULT 6
US-09-949-016-13002/C
Sequence 13002, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241, 755
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 60/237, 768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231, 498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastaSBQ For Windows Version 4.0
SEQ ID NO: 13002
LENGTH: 168394
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) .. (168394)
OTHER INFORMATION: n = A, T, C or G
US-09-949-016-13002

Query Match 34 5*: Score 2615.4; DB 4; Length 168394;
Best Local Similarity 85.1%; Pred. No. 0;
Matches 3133; Conservative 134; Mismatches 355; Indels 58; Gaps 33;

Qy 1012 TTCCCAATTATGACGCCCTCCAAAGCAGTGGAGGAGAAATGGCCAGCCAGTGC 1071
Db 29005 TCCCCAACGACAAACACCCCTAAATTTCTGGAGAATTGGCTGTGAGTC 28946
Qy 1072 CATGGCYTTTYYCTCCAGACTTAACCGAAATAAACAGACTTAGTTAACTCTCA 1131
Db 28945 TATGTTACCTTTCCCTGTAGACTTGAAAGCAATTAAATGACCTAGTTAACTCTCA 28886
Qy 1132 GATAAATCTGGCTATATGTTCAAGGTTAGCAACATTTGTGACA 1191
Db 28885 GATAACCTGTTGGCTATATGATGCTTACAGGGTGAACGAACTTC 28826
Qy 1192 TGGAGAGATATATGTCATGCTAAATCAGACATAACCCAAATGAGAGAGTCAC 1251
Db 28825 TGGAGAGATATATGTCATGCTAAATCAGACATAACCCAAATGAGAGAGTCAC 28767
Db 1252 CATAACTGCGCCTGAGRTTGGCATCTGTTGATCTGAGTCATGGATANG 1311
Db 28766 CATAACTGCGCCTGAGATTTGGTGTGATCTGTCAGTCAGTCA- ATAGG 28709
Qy 1312 ATGACACAGACAGAAATCTGAACTGGAGATGGTGTGAGAGATCTCTA 1371
Db 28708 ATGACACAGACAGAAATCTGAACTGGAGATGGTGTGAGAGATCTCTA 28650
Qy 1372 CCCTCAATGCGACACAGAAATCTGAACTGGAGATGGTGTGAGAGATCTCTA 1431
Db 28649 CCCTCAATGCGACACAGAAATCTGAACTGGAGATGGTGTGAGAGATCTCTA 28595
Qy 1432 CTGGTGTGCTASAAGGAACTAAGGAAARTCTAYGAATTACTCAATGATG 1491

Db 2765 AGAATAGCCATCATCTATGCCCTGAGACACAGGGGACAATGTCGGATA 2824
Qy 7478 AACCCAAGTTTCGAGGCCAACGGCACACCCCTTG67CCCTCCCTTGATGCA 7537
Db 2825 AACCAAGTTTCGAGGCCAACGGCACACCCCTTG67CCCTCCCTTGATGCA 2884
Qy 7538 GCTCTGTTTCAGCTTACTCTTACTCTTAACTCTGCACTGCR 7522
Db 2885 GCTCTGTTTCATCTTACTCTTAACTCTGCACTGCA 2929

Db 28477 GGATTTGGGAGC-AACTCTCTGTAACCTGACTCTGACTCTGTAAGGGC 28419
Qy 1612 GCGTAAGTTTACTCTGCACTGCAAGCATAG AAAAATCTCAAAGTCTGGCGT 1670
Db 28418 GGTAAAGTTTACTCTGCACTGCAAGCATAG AAAAATCTCAAAGTCTGGCGT 28359
Qy 1671 AGSCCGGGAGCAACTTAACTGAAACCTTATGAACTTGCACACTTCGTTTATAG 1730
Db 28358 AGSCCGGGAGCAACTTAACTGAAACCTTATGAACTTGCACACTTCGTTTATAG 28299
Qy 1731 AGTCAAGGGAGCAAGGGAAAGGAAACGAACTTAAAGAAGGACACCGGTTAG 1790
Db 28298 AGATCAGGAGGAGCAGCAACGAGTAAAGGAA-TAAAGAAGGACACCTTAG 28240
Qy 1791 TCATGACCTCAGGCAAG-TGGACTTGGAGETCTGAACTGAAAGGAAAGCTGGAAAT 1849
Db 28239 TCAATGCGCTCAGGCAACGACTTCGAGGCTCTGAGACAGGAAAGCTGGAAAT 28180
Qy 1850 TGAATGCTTAAATGGCTGCTCCAGTCAGTCGGGTCAAGGAACTTTAAAAGATGT 1909
Db 28179 TGAATGCTTAAATGGCTGCTCCAGTCAGTCGGGTCAAGGAACTTTAAAAGATGT 28120
Qy 1910 CCAAGTGAAGTAAAGC-GCCCTTGTGTCATGCCCTTATTCAAGGAAATCTGGAA 1968
Db 28119 TCTATAGATAAGCACCCTCTGTCATGCCCTATGTCAGGGAAATCTGGAA 28060
Qy 1969 GGCCCACTCCCGAGGGCAACGGTCTTGACTGAGCCACTAACCGAGATGATCCA 2028
Db 28059 GGCCCACTCCTCCCGAGGGAGCAAGGTTCTGTAGAGCCACTAACCGAGATCCA 28000
Db 2029 GCAGCAGGAGCTGAGGGTCTGGGCAAGGGCATCCATGCCATCACCTCACGGCC 2088
Qy 27999 GCAGGAGACTGAGGGTCTGGGCAAGGGCCATGCCATCACCTCACGGCC 27940
Db 20899 CTGGGTACTTGGCATTGAGGAGGAGGT--TGTCTCTGACACTGGTGGGT 2145
Qy 27939 CCGGGTATGCTTAACCTTGAGGGCCAGGGTAACTCTCTGAGACTGGGGGG 27880
Db 2146 CTCTCTAGCTTACTCTCTGCGGAGAACCTCTCTGAGACTCTGTCCTATCTGAG 2205
Db 27879 CTTCCTGAGCTTACTCTCTGTCCTGTCAGACTGT-CTCAGATCTGTCACTATCCAGG 27821
Qy 2206 GGGGCTNTAAAGCAGGGCACTCTAGATACTTTCAGCCACTAAGTATGACTGG 2265
Db 27820 GG-----TCTGGAGACACTAAGAACCTCTGAGACACTAAGCTGTCG- ACTGG 27770
Qy 2266 GGAGCTTATCTCTTACATGCTTCTCTAATATGCTGAAGGCCACTACTTGTAA 2325
Db 27769 GGAACCTTACTCTTACATGCTTCTCTAATATGCTTAACCCACTCCCTGTTA 27710
Qy 2326 GGGAGAGACATTCTAGAAAGCAGGGCATTATAACCTGAACTAGGAAGGACA 2385
Db 27709 GGGAAAGACATCTCAAAAGACAGGACATTATACTGAACTAGGAAGGACA 27650
Qy 2386 CCCGTGTGTTGTCAGGAGGATTAATCTGAACTGTCGCAACAGGAG 2445
Db 27649 CCCTGTTGTTGTTGAAAGAACTTAACTCTAAAGTCTGGGCAACAGGAG 27591
Qy 2446 ACAATGAGAGACAAAGATGCCGTCCTGTTGAGTAACTAAGGATTCACCT 2505
Db 27590 ACAATGAGAG- CAATATGCCGTCCTGTTGAGTAACTAAGGATTCACCT 27532
Qy 2506 CCTTCCTTACAAAGGAGTACCCCTGAGCCAAAGGAGTCACAGAAGA 2565

QY	27531	CCTTCCCTACCAAGGAGTAGCCCCCTAGACCGAGGCCAACAGACTCCAAA	27472
QY	2566	TGTGTAAGCTAAAGGCCPMAAGCTTAGTAALACCAGTCATRACTCCCTGCGTAATT	2625
Db	27471	TGTGTAAGGACTAAAGCCCTACTCTCTAGTAACCATAGCCCTGCACTC	27412
QY	2626	CGTAGTGAGATTAGGGAGCACAGAAACCGAGCTGAGCACTGGAGGTTAGTGCAAGATCT	2685
Db	27411	C-----AATTAGGAGTAAAGAGGCCAATGAGACTGGA-GGTTAGTACAGATCT	27360
QY	2686	CAGGATTATCAATGGAGGCCGTTCTTATACCCAGCTGAGCTAGCCCTATCTG	27355
Db	27359	CAGGATPATCACT-GAGCTGCTGTTCTATACCCAGCTGAGCTACCTTACTC	27301
QY	2746	TGMYTCCAAATACCGAGGAGCAGAGCTGAGGTGTTACASTCCAGCTGAGCTAGCCCTATCTG	27305
Db	27300	TGCTTCCCAAATACCGAGGAGCAGAGCTGAGGTGTTACAGTCTGAGCTTACAGTGCCT	27241
QY	2806	TCTTCTGCATCCTGTACATCCGACTCTAATCTCTTCTTCTGAGAGTACTICA	2865
Db	27240	TPTCTGCATCCTGTACATCTGAGCTTACCTCAATCTTCTTCTGAGAGTACTICA	27181
QY	2866	ACCCACATCTCACTACTGGACTTTACCCAAAGGGTCAGGATAGYCCCCTAC	2925
Db	27180	ACCCAGTCTCACTACTGGACTTTACCCAAAGGGTCAGGATAGYCCCCTAC	27121
QY	2926	TATTOGCCAGGCTATAGGCCAAGCTGAGCTTGAGYCARTNTCATCTGAGCTCTGRC	2985
Db	27120	TATTCTGCCAGGCTATAGCTCAGACTTGAGAGCTGAGCTCTCATGACGTTGRC	27061
QY	2986	TTCRSTAKGKGATGTTACTTCTCYCCYRT-CAGAAGCTGCCCCATGAGCC	3044
Db	27060	TTCRSTAKGKGATGTTACTTCTCYCCYRT-CAGAAGCTGCCCCATGAGCC	27001
QY	3045	ACCCAGCRCTTMAATTCTCGCVACTCTGCTACAWGCTTCCAACCSAROGCT	3104
QY	3165	CTCTGTGAGBAYRATCCACCTTACTGGTTATCTCATCYAAACCTTAAGCA	3224
Db	27000	ACCCAGCCTTAACTCTCACCTTGCTCAGAAGTTCCAACCAAGCT	26941
QY	3105	CARCTCTGCTCACAGGGTTAACTTCTGCRATAATTCCAAAGGACARGCC	3164
Db	26940	CAGCTCTGCTTACAGCAGGTTAACTCTAGGCTTAAATTCCAAGGACAGGGCC	26881
QY	3225	CTAAGGRRTCTCTGGCRTAYAGGGTCTGCCGAATATGGATTCAGGCTTAACTCCATCAACCA	3284
Db	26880	CTCACTGAGAATGATCCAGCTTACTGGCTTAACTCCATCAACCA	26821
QY	3285	AATACCCAGGYCATHAWATACASTATTAAAGAACCTCAGAAGCCATACCTTART	3344
Db	26762	AATACCCAGGCAATTATATACACTTAACTGAGCTGAGAAGCCATACCTTART	26703
QY	3345	AGATGTGAYAMTGTAGGAGMRAAGTGGTTCCAGGCCCTAAAGAAGGCCCTAAACCAA	3404
Db	26702	AAATAGGACTCTAAAGCAAAAGCAGCTTCAGGCCCTAAAGGCTTAAACCAA	26644
QY	3405	GYCCAGTGTAGGACTTACACAGGGCAAGCTTCTCATGTCATGTCAC--AACCAA	3464
Db	26643	GACTGAGTGTAGGACTTACACAGGGCAAGCTTCTCATGTCATGTCAC--AACCAA	26587
QY	3465	AGRAVAGCTCTRGAGTCTTACAGCAAGTCRCCAGGAGACCTGCAACYRTGRIA	3524
Db	26586	AGAATAGCTCCAGGAGTCCTACAGGTTCAAGGAGCTGCAACCATGCCATA	26527
QY	3525	CCTGASTAAAGAAATGATGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT	3584
Db	26526	CCTAAGTAAAGAAATGATGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT	26467
QY	3585	GGCAGTAGCTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT	3644
Db	26466	GGCAGTAGCTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT	26407
QY	3645	GTGGACATCTATGKGTAvgRCATCTGCTAAAGGAGCTGTCGTCGAG 3704	
Db	26406	GTGGACATCTATGKGTAvgRCATCTGCTAAAGGAGCTGTCGTCGAG 26347	
QY	3705	CAACGTTCTTAATRAGGCTTACTGAGCTTACAGCTGCTAAAGGAGCTGTCGAG 3764	
Db	26346	CAATCGTTCTTAATRAGGCTTACTGAGCTTACAGCTGCTAAAGGAGCTGTCGAG 26287	
QY	3765	TGTCACACTTAACCCAGCNATCTCCAGACATGAGAGAGATRAYATA 3824	
Db	26286	TTGCCAACTTAAACCCCACTATTCTCCAAACAGA - - AAAAGATAAACATAA 26229	
QY	3885	GACTGATCCYACCTTAACCTGTTACTGCTGAGTAACTGAGTACGCTC 3944	
Db	26169	GACTAATCC-TCACTTGTAAATTATGAAATTCTTGTAAAAAAGGACTTC 26111	
QY	3945	GAAGAGGGTATCAGTGTACCTTACCTGTTAGTGAAGAAAGGACTTC 4004	
Db	26110	-AAAAGCAGGTTAGCTGAGTGTAGTAATATGGATACTGTAAGTAACCCCTACTC 26052	
QY	4005	CAGGACTATGTCAGTGTAGGCTACTGAGTATGGATAATTGAACTTCTGCTC 4064	
Db	26051	CAGGACTCTGGCTGCGAGCTTACAGCTTACAGCTTACCTGCTCTAGGCC 25992	
QY	4065	AAGRAAAAGGYYAATATATACAGACTCTRARTATGTTGAGTCTTGTAGAAGTCT 4124	
Db	25991	AAGGAAAMAGGTTAAATACATACAGACTCTAAGTATTCTTACCTGCTC 25932	
QY	4125	ATGNGCAATATSAGAGAAGGGATTCTACTCTYGGRRGAACCTTCACMACAT 4184	
Db	25931	ACAGCAATATGGAAAGGAAATCTACTCTGAGGAACTTCACCTCAACATC 25872	
QY	4185	AGGAGCCATTAGGARATTATXGGWGTACAGAACTTACAGGTTGAGTCT 4244	
Db	25871	AGGAGCCATTAGGAGTTATATGGCTGAGAACCTTAAGGAGCTGAGCTTAC 25812	
QY	4245	ACTGCGGGTCACTCANAGGAAGAGAAGRAAGGAAATASAGRGAYTGCAGKATA 4304	
Db	25811	ACTGCGGGTCACTCANAGGAAGAGAAGRAAGGAAATASAGRGAYTGCAGKATA 25752	
QY	4305	TTGAGCMAAAGAGCTGCAAGGGAGCCCTCATTAGAATGTTATTAACTCCCT 4364	
Db	25751	TTAAAGCTTAAAGGCCAACAGGAGAACTCTCCATTAAAGCTTATAGAAGGACCC 25692	
QY	4365	TAGTATGGTAACTCCCTCGGGAACCAAGCCGACTCTAGCAGGAGAAAGAAT 4424	
Db	25691	TAGTATGGTAACTCCCTCGGGAACCAAGCCGACTCTAGCAGGAGAAATAGAC 25634	
QY	4425	GGGGACCTCAG---GAGGAGTTCTCCCTCGGAGCGGTAGCCACTGAGAGGA 4481	
Db	25633	GGGGACCTCAG---GAGGAGTTCTCCCTCGGAGCGGTAGCCACTGAGAGGA 25574	
QY	4482	AAATACTTCTGCTGCAACTTCACTGAACTTAAACCTTCATCAACCTTC 4541	
Db	25573	AAATACTTCTGCTGCAACTTAAACCTGAACTTAAACCTTCATCAACCTTC 25514	
QY	4542	ACTTAGGSCATCGTAGCACCCCATCARATGCCAACTTATTACTGAGGCCCTT 4601	
Db	25513	ACTTAGGCTTAACTGACCCATGAGCTGCAACATATTACTGAGGCCCTT 25454	
QY	4602	TCAAACTTATGAGGAACTTCACTGAACTTAAACCTTCATCAACCTTC 4661	
Db	25553	TCAAACTTATGAGGAACTTCACTGAACTTAAACCTTCATCAACCTTC 25394	
QY	4662	CTYATCCAACTGCTC 4681	
Db	25393	ACTGCAAGCCCATCATICA 25374	

Qy 5652 ATGCCGCTGATGACCAGTAGTCCTCCYACCMAGAGTTCTATGGAGATAAGCAGGGTCC 5711
 Db 834 ATGCCGCTGATGACCAGTAGTCCTCCYACCMAGAGTTCTATGGAGATAAGCAGGGTCC 893
 Qy 5712 CGGAATATTGATGCCCATGTTAGAGCTTSTAAGGAACCCACTTCAC 5711
 Db 894 CGGAATATTGATGCCCATGTTAGAGCTTSTAAGGAACCCACTTCAC 953
 Qy 5772 CCACACCAATATGCCGCACACTGTCTACTCTGCCACTCTTGCA 5831
 Db 954 CCACACCAATATGCCGCACACTGTCTACTCTGCCACTCTTGCA 1013
 Qy 5892 CTGTTGACTTACTCACCAACTGTGTTCTGAGTGGGGTAGTCA 551
 Db 5832 TCATTTGACGAGAAATGATTAATCTTAGTGTCTGGAGACTGAGTC 5891.
 Qy 1014 TCATTTGACGAGAAATGATTAATCTTAGTGTCTGGAGACTGAGTC 1073
 Qy 5952 AAGAGAAACATGTAAGAGAATCTCCAACCTCACCGGGTACATG 6011
 Db 1134 AAGGAANAAATGTAAGAGAATCTCCAACCTCACCGGGTACATG 1133
 Qy 1074 CTGTTGACTTACTCACCAACTGTGTTCTGAGTGGGGTAGTCA 5951
 Db 6912 GATCCCTGATCTCCCTCTAGAACCTCTAGGCTATATGCTACTCC 6971
 Qy 2094 GATGCCCTGATCTCCCTCTAGAACCTCTAGGCTATATGCTACTCC 2093
 Qy 6972 TGGACCTGATCTTTRACCTCTGTACTTGTCTCTCCAGAATGAGT 7031
 Db 2154 TGGACCCCTGATCTTTRACCTCTGTACTTGTCTCTCCAGAATGAGT 2213
 Qy 7032 ACTACAATGGAGGCCAGATCGACTAACCTGAGTCTGGCC 7091
 Db 2094 ACTACAATGGAGGCCAGATCGACTAACCTGAGTCTGGCC 2213
 Qy 6072 GTTAAACCTTTATACCCCTACTGGCTCATGGTCTGGCCAAACCTC 6131
 Db 1254 GGTAAAGCCTTTATACCCCTACTGGCTCATGGTCTGGCCAAACCTC 1313
 Db 1194 CCCCTACAAAGGAGCTAGATCTCAAACTACATGAACCTCTGCT 1253
 Qy 6132 TAATGTTGATATGCCCTGACTCTGRCATATGTTCAATCTGACTGA 6191
 Db 1314 TRACTTGTGAAATRACCAACTCTGACTCTGAGCTTGTAAATTGCA 1373
 Qy 6252 TTCCATCTGAAATPACCAATACCTCAAACTCACCTGAAATGCT 6311
 Db 1434 TTCCATCTGAAATPACCAATACCTCAAACTCACCTGAAATGCT 1493
 Qy 6312 ATACACACCAACTCCAACTGATCAGGTGGTAACTCTCCAC 6371
 Db 1494 ATACACACCAACTCCAACTGATCAGGTGGTAACTCTCCAC 1553
 Qy 6372 ACCCTCAGGATATTTCTCTGGTACTCTGGCTACATGTTGATGCTC 6431
 Db 1554 ACCCTCAGGATATTTCTCTGGTACTCTGGCTACATGTTGATGCTC 1613
 Qy 6432 AGATCTATGCTCTCTCATCTAGGCCCCATTCACCTGACAAGA 6491
 Db 1614 AGATCTATGCTCTCTCATCTAGGCCCCATTCACCTGACAAGA 1673
 Qy 6492 TTATACAGTTATGTCATATAAGCCCCAACAGTACCATTCCTTGT 6551
 Db 1674 TTATACAGTTATGTCATATAAGCCCCAACAGTACCATTCCTTGT 1733
 Qy 6552 TATAGAGCGGAGGTCTAGTGCACTAGTCATGGCTATCAC 6611
 Db 1734 TATAGAGCGGAGGTCTAGTGCACTAGTCATGGCTATCAC 1793
 Qy 6612 TCACTCTACTACAACCTCTCAAGACTTAATGGGAATGGAACGG 6671
 Db 1794 TCACTCTACTACAACCTCTCAAGACTTAATGGGAATGGAACGG 1853
 Qy 6672 OCTGGTCACTTGCAGATGACTTAACCTCTAGGAGCTGCTTAAATG 6731
 Db 1854 CCTGGTCACTTGCAGATGACTTAACCTCTAGGAGCTGCTTAAATG 1913
 Qy 6732 AGCTTAGACTYCTAACCGTCTGAGAGGGAAACCTGTTT 6791

RESULT 9
 US-09-573-080A-26
 ; Sequence 25 Application US/09573080A
 ; Patent No. 682897
 ; GENERAL INFORMATION:
 ; APPLICANT: JOAN, KNOELL
 ; APPLICANT: ROGAN, PETER
 ; FILE REFERENCE: 30307
 ; CURRENT APPLICATION NUMBER: US/09/573, 080A
 ; CURRENT FILING DATE: 2000-05-16
 ; NUMBER OF SEQ ID NOS: 479
 ; SOFTWARE: Patentin version 3.0

SEQ ID NO 25
 LENGTH: 8399
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: repeat region
 LOCATION: (1) .. (8399)
 OTHER INFORMATION: hervg
 PUBLICATION INFORMATION:
 AUTHORS: Jurka, J.; Wilchiewicz, J.; Milosavljevic, A.
 TITLE: Prototypic sequences for human repetitive DNA
 JOURNAL: Journal of Molecular Evolution
 VOLUME: 35
 ISSUE: 4
 PAGES: 286-291
 DATABASE ENTRY NUMBER: Database of repetitive elements (repbase)
 DATABASE ENTRY DATE: 1996-01-26
 US-09-573-080A-26

Query Match 28.6%; Score 2168; DB 4; Length 8399;
 Best local similarity 67.6%; Pred. No. 0; Mismatches 1354; Indels 103; Gaps 34;
 Matches 3334; Conservative 143; Mismatches 1354; Indels 103; Gaps 34;

```

Qy      605 TGATGGAAAGTTCGCCGCAAGACAAAGCCCTTAAGAGTATTCTGAGTGCCTG 1344
Qy      1276 TGTAGGAACACTCGGCATCACAGCCTAACCTTGAAATGATCTGACCTGG 1344
Db      665 AMCAATTGACCTCTAGACATAAGAAAGACACTATTAATTCTCTGAGTGCCTG 724
Db      1336 ACCAAATTGACCAACACCTGAAAGAGGGCAGCTCATTTTCTGACCTACAGCT 1335
Qy      725 GGCAC-----CTGAGGGAGATAAATATAAAT 752
Db      1396 GACCTPAATACTCTCTCTCTGATGGGAAMATAAGCCACCTGAGGGAGATAATTACA 1455
Qy      753 ACACCATCTACAGCTAGACVCTTTGT--AGAAAGCAATGAGTGAGTGCAT 809
Db      1456 ATACATCCCTCAGGTTGACCTTTCTGTAAGAGGGAGTAATGAGGTAATACAT 1515
Qy      810 AAG----TACRACTTCTTCATTAAGAGACAATCACATATGTAAGAGTGTAT 865
Db      1516 AAGGTTTACAGCTTCTTCTGAGGAGAAATCACAACTATGCAAGCTGCAAT 1575
Qy      866 TTATGCCCTCAGGAAAGCCTCAGACTCTCCCTATCCAGATCCGACTCTC 925
Db      1576 TTACATCCCAAGGAGGACCTTCAGCTTACCCATATCTTAGCTCCCTAGACTCTC 1635
Qy      926 C---CCAMYTAATAAGGACCCTGAACTAAATGTCAAAGGAGATAGACAAAGG 983
Db      1636 CTCTCATATAGGCAAGCCCTCCATTCATGCGAAAGAATAGACAAAGA 1695
Qy      984 GTAAACAGTGTACCAAGAGTGGCCAAATATGCCCAATATGACAGCTG 1043
Db      1696 ATCTCCAAAGGGRCAACAAACCCCTGGGATATCGGTATGTCCTT-GAGCTGTAGG 1754
Qy      1044 AGGAAGAGAGTGGCCCAAGGAGCTGCTGTGCTTTCAGCTAACCTAAAGCA 1103
Db      1755 GGTTGAGAAGTGGCCAAACCCAGCTACATGTCCTCTGCTGTTAAGCA 1814
Qy      1104 ATAATAACAGACTTAGTAAATCTCAGATAAYCTGTGATGCTGATATTGTTACA 1163
Db      1815 GATCAGGAGGACTGGGGAGTTCTAGATGATCTGATGAGGACATAGATGCTCA 1874
Qy      1164 AGGTTAGGAACTTCTTCTGACATGAGGATATATGTCACTOCTAATCAGA 1223
Db      1875 GGGTCTAGGGCAACCTCGATCTGACTTGAGGAGATG-CATGTTATGTTAGCTAA 1933
Qy      1224 CACTAACCCAAATGAGGAGTGGCCACATRACTGAGCTGAGGTTGGGATCTC 1283
Db      1934 CCCTGGCTTAATGAAAGAATGTTGGCTTAGCTGAGCTGAGTTGGAGATACTA 1993

```

Qy 1284 GGATATCTCACTCAGGTCAATGGATANGGAGACACAGAGGAAAGAAATGATCCCCA 1343
Db 1994 GGTATCTTAGTCAGTAATG--ATAGAATGACAGCGAGAGATAATTCCTA 2050
Qy 1344 CAGGCCAGCARGCAGTCAGCTTASACCTCATGGAGCTGCTAAGAGAAATGACATG 1403
Db 2051 CTGGCCAGCAACCCAGTATGATCTGATCCCACACTGGACCT 2106
Qy 1404 GGAGATGTTGCTGAGCATTTGCTACTTGCTGCTAAGAGACTAAGTGGAAACTSG 1463
Db 2107 GGGAACTGGAGTCGTAACATCTGAGTGTGTTCTGAAAGGACTTAACTG 2166
Qy 1464 AAGAAARTCTAYGATTACTCAATGATGTCACCTAACAGGGAGGGAGAAATC 1523
Db 2167 AA-AAAAGCCATGATTATCAATGATGTCACCAATAAATCAAGGAAAGAAATC 2224
Qy 1524 ACTCTCTGCTTCTGAGAGACTAAGGAGCATGGAGCTTAAAGGAAAT 1583
Db 2225 CTTCGCTCTTCAGGGCTTCAGGAGCTTAAAGGAAAT 2284
Qy 1584 ACTCTCTGCTTCTGAGAGACTAATCTTAAGGCTAAGTGTATCACTCAGTCAGTCAGTCAGCA 1613
Db 2285 AATCACTCGAGGGTCATGATCTTAAGAGATAATGTTATACCCATCAGCAGCA 2344
Qy 1644 TTAG-AAAALACTCAAAGTCTCGGAGGCCCCGGAGAACTTAGAALACCTTATGA 1702
Db 2345 TCAGPAGAAAGCTCCAAAGCAAGCAGCCCTGGCCCTGAGCAAAATATGAGGCAATTAA 2404
Qy 1703 ACTGGCAACYCTGGTTTATAATAAGATGAGTCAAGGAGGAGGGAGGAGGAGAAC 1762
Db 2405 ACCCTGGCAACCGGGTGGTCTATAATAGGACCAAGGAGAACGGCCAATAGGAA 2460
Qy 1763 GGGATTAAAGAAGGCCACCGCTTTAGCTGACCTTCACTGGCAAGGAGGAGGAGG 1822
Db 2461 AGAGATCAGAAAGGCGCAGCTTATGCTAGGGCTTAGACAACTTGGT 2520
Qy 1823 TCTGGAAGGAAAGCTGGCAATAGTGAATGCTTAATAGGTTGCTTCACTGGGT 1882
Db 2521 TCAGGAGGACAGAAATGAGGAGGCAATCAGCTGGTGGGGCTTGTATCAGTGTT 2580
Qy 1883 CTACAGGAGACTTAAGAAGTTGCAAGTAGAGTAAGCCCTTGGCATGC 1942
Db 2581 TTACAGGAGACTTAAAGATGTCAGTGAAGACAGCTGGCCCTTGTGATGC 2002
Qy 1943 CCCTTATTCAGGAGACTCTGGAGACGCCACTGGCCAGGGACAAAGGCTTGTG 2640
Db 2641 CTACTGGGAGGACATCTGGAGGAGTGGTCACTGCCCAGAGGACAACTTCTG 2700
Qy 2003 TCAGAGCCACTAACAGATGTCAGGAGCAGACTGAGGAGCTGGCTGGCAAGGGCCA 2062
Db 2701 TTGGAGGCCAAACCTGTGATTCACACAGGACTGAGGAGTCCACGGCAACGCC 2760
Qy 2063 TCCCTGGCATACCTCTGGAGGCCAGCGTATGCTGCTGAGGAGGAGGAGA 2119
Db 2761 GTCTCATGTCATCACCTCTACTGAGCCAGGATGCTGAGGAGGAGGAGGAGA 2820
Qy 2120 GGTGTTCTCTGGACATCTGGGGCTCTCTAGTCTACTCTCTGTCCTGGAGA 2179
Db 2821 GACTCTCTGGACATCTGGAGCTGAGCTGCTCTGCTGCTTCTGTTGATGTC 2880
Qy 2180 TCCCTGGCATCTGTCATATTCTGGGGGTCTTAAAGGGGAGCTGATCTT 2239
Db 2881 TCCTCAAGCTCTGTTACATCC--GAGGAATCTGAGGCTGAGCTGTT 2936
Qy 2240 YTCCGAGGACTAGTATGAATGGGGAGCTTATCTCTGCTGCTTCTGTT 2999
Db 2937 CTCCCAACCTCTGTTGTT-AATGGAGACTTGTCTTCTGTT 2995
Qy 2300 TGCTGAAAGCCCAACTACTCTGTTGGAGAGCATCTAGCAGAGGGCCATTA 2359
Db 2996 TGCTGAAGTCCACCTTATGAGGAGTAACTT-TAGGAAAGTGGAGCTTA 3053

Db	354	AATACCCATTGCTGT-CCCCTGCTTGGAGGA-TAACTGAGTCGGCATAG	412
Qy	2442	AAGGAAATATGGAGCAGGCCAAGAATGCCGCTCTCACTTAACTAAGGTTCC	2501
Db	413	AAGGAAATATGGAGCAG-CAAGAATGCCGCTCTCACTTAAGGATCT	471
Qy	2502	ACTCTCTTCCCTACCAAAGGCATACCCCTCAGGCCAAGGCCAACAGGATCCA	2561
Db	472	GCTCTCTTCCCTACCAAGGCTTAGAAGCCGCTCACAGGAGTACCTCTAGACC	531
Qy	2562	AGATGTTAAGGACTAAAGCCAGGCTTAGTAAACCTATGCTAACMCCCTGAGT	2621
Db	532	AGATGTTAAGGACTAAAGCCAGGCTTAGTAAACCTATGCTAACMCCCTGAGT	591
Qy	2622	AATCCGTTAGGATGAGGAGGACAGAACCCAGTGACAGTGGAGGGTAGCAG	2681
Db	592	ACGCC-----AATTAGGAGTAAGAAACCAGGAGCTGAA-GTTAGTCAG	643
Qy	2682	ATCTAGGATTCATGGAGGCGTGGCTTAACTACCGCTTACCTAGCCCTAT	2741
Db	644	ATCTAGGATTCATGGAGGCGTGGCTTAACTACCGCTTACCTAGCCCTAT	702
Qy	2742	ACTGTGMYTCCCAAATACCAAGGGAGCAGAGTGT'TACSTCTTGACCTT	2801
Db	703	ACTGTGMYTCCCAAATACCAAGGGAGCAGAGTGT'TACSTCTTGACCTT	762
Qy	2802	GCCTTCTCTGATCCCTGTACTCTGACTCTCACTCTGTTTSCCTTGAGA	2861
Db	763	GCCTTCTCTGATCCCTGTACTCTGACTCTGACTCTCACTCTGTTTSCCT	822
Qy	2862	TCAAACCCARCATCTRACTCACCTGGACTRTTACCCAGGGTCAGGATCCCC	2921
Db	823	TGAAACCCARCATCTRACTCACCTGGACTRTTACCCAGGGTCAGGATCCCC	882
Qy	2922	CATCTATTGGCCAGGATAGGCCAACACTCTGAGCTGTTACCCAGGGTCAGGAT	2981
Db	883	CATCTATTGGCCAGGATAGGCCAACACTCTGAGCTGTTACCCAGGGTCAGGAT	942
Qy	2982	GTCTCTCRATAKGTTGATGATGTTTTRGCGCCTACAGGCTTACCTGAGCTT	3041
Db	943	GTCTCTAGTACGGAGTACTTTAGTGCCTTACCTGAGCTTACCTGAGCTT	1002
Qy	3042	GCCRCCCAAGRCCTTMAATTCTCCYACCTGTSCTACWGTTCCAAACRAR	3101
Db	1003	GCCRCCCAAGRCCTTMAATTCTCCYACCTGTSCTACWGTTCCAAACRAR	1062
Qy	3102	GCTCTCTGTCACAGGGTTAACTTGTGCTTACCTGAGCTTACCTGAGCTT	3161
Db	1063	GCTGGCTCTGTCACAGGGTTAACTTGTGCTTACCTGAGCTTACCTGAGCTT	1122
Qy	3162	GCCCTCAGTGGAGGARYATTCAGGCTTACAGGGTTAACTTGTGCTTACCTG	3221
Db	1123	GCCCTCAGTGGAGGTTAACTTGTGCTTACCTGAGCTTACCTGAGCTTACCTG	1182
Qy	3222	CAACTAAGGRRCTCTGCTTAYASGTTCTGGCGAWATGATGCCGAGGTAG	3281
Db	1183	CAACTAAGGRRCTCTGCTTAYASGTTCTGGCGAWATGATGCCGAGGTAG	1241
Qy	3282	CAAAATGGCAGGATTWATCSTAATTAAGGAATCTAGAAGCAATACCTATT	3341
Db	1242	CCCATAGGCCAGACCTATPACTCTAATTAGGAACTCTAGAAGCAATACCTATT	1301
Qy	3342	ARTAGATGGAYAMCTGAAAGYMRAIGGGTTCAGGCCCTAAGAAGGCCCTAAC	3401
Db	1302	AGTAGATGGACACTT--ACAGAATGCTTCCAGGCCCTAAGAAGGCCCT-AAC	1356
Qy	3402	CAAGGCCAGTGTAGYTTGCAACRGSGCAAGCTTTCAGGCCCTAACGAAAG	3461
Db	1357	CAAGGCCAGTGTAGYTTGCAACRGSGCAAGCTTTCAGGCCCTAACGAAAG	1415
Qy	3462	NACGAAAGCTCRRGGACTTACAGCGTTCAGGAGYAGCTGCAACYRGCG	3521

PRIOR FILING DATE: 2000-10-03
 PRIORITY APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SEQ ID NO: 13840
 LENGTH: 276687
 TYPE: DNA
 ORGANISM: Human
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1..(276687))
 OTHER INFORMATION: n = A,T,C or G
 US-0-9-949-016-13840

Query Match Local Similarity Score 1507.4; DB 4; length 276687;
 Matches 22891; Conservative 68.3%; Pred. No. 0; Mismatches 914; Indels 52; Gaps 24;

QY 697 CGACTTATTCCTCTGCGAGTGCGCCTCGCACTCTGAGGGAAATAATTAAACAC 1591 TGAGGCCAACTAACTTAAGGTAAGTTTACACTGAGCTGAGACATTAGAAA 1650 QY 28073 TGGGGCAATGATGTTAAAGATAAGTTTACACCGATCCAGCCAGATTCAGGG 28072 Db 27184 CCAAATTTCTCTGATGGGAAAATGCCACCTGAGGGAACTACATAC 1651 AAACTTCAAAGCTGCGTAGGCCCGAGCNAACTAGAACCTTATGAACTTGGCA 1710 QY 27243 TATCTCGAGCTTGACCTTCTTAAGAGGAATGGGAACTGAGGAACTACATAC 1711 ACYTCGTTTTATAGAGATCAGAGGAGCAGGGAAAGGAAACTTAAACCTGAGCA 1770 Db 27244 TATCTCGAGCTTGACCTTCTTAAGAGGAATGGGAACTGAGGAACTACATAC 1770 QY 27303 ACCMCGGGTTCTATAAGGACATGGGATGTTACCTTGT 28193 AAAAAGGCCACGCTTACTGAGCTCAGGCAAGTGGCTCTGAA 1830 Db 27304 CCAAGCTTCTTTCATGGGAAATCACCTACAATTTGAAAGTGTGTTATGCC 1831 AGGGAAAGCTGGCAAATGATGATGCCATAAGGCTGCTGCTTCCAGTCAGG 1890 QY 27364 CACAGGAGGACCTCTCAGTCACCCCATACCTGCTCCATAGTCCTCCAT 1891 ACRCTTAAAGATGTTCCAGTAGAGCTAGAAGTGGCTGCTCCAGTCAGG 1890 Db 27364 CACAGGAGGACCTCTCAGTCACCCCATACCTGCTCCATAGTCCTCCAT 1891 QY 27423 YTTATAAGGACCCCTCAACCAAATGTCACAAAGGAGATGACAAAGGTTAA 27363 QY 27424 GAACTGATATCTCTCTATCTCCCCGGCCAGAAGGAAATCTCA 27423 1951 TCAAGGAAATGGGAGGCCAACTCACCTGGTGTAGGCTGTTATCAGTGTTACAGG 28371 Db 27424 GAACTGATATCTCTCTATCTCCCCGGCCAGAAGGAAATCTCA 27423 QY 27432 CCCAAACAGTATGATCCACACAGGCTGAGGTCTGAGGCTGCTCCCTGTG 28432 CGAGGGRATACTGGAGGAGTGCACTCGCCAGAGGAATTCCTGGGTCAGAGC 28432 Db 27432 CCCAAACAGTATGATCCACACAGGCTGAGGTCTGAGGCTGCTCCCTGTG 28432 QY 27441 GTGAAACCAAGAGGAGGCCAATATTCCAATATGACCCCTCCAGAGTGGAGA 2011 CACTAACAGATGATCAGCAGCAGGAGCTGAGGGCTGGGGCAAGGGCATCCATC 2011 QY 27441 GTGAAACCAAGAGGAGGCCAATATTCCAATATGACCCCTCCAGAGTGGAGA 2011 Db 27442 AAGGACCAAAACCTCTGGCTATTGCTT-TCAAGTGAGGGAGG 27442 2071 CATCACCTCATAGAGCTGGTATCTGACCATGAGGGCAGGA- -GGTTCT 2127 Db 27442 AAGGACCAAAACCTCTGGCTATTGCTT-TCAAGTGAGGGAGG 27442 QY 27452 CATCACCTCATAGAGCCGGGTATTTAACATATGAGGGCCAGGAATTGACTCT 28552 CTCAGGAAATGGGAGCTTCTGAGGGGGCTGGGTATTTAACATATGAGGGCCAGGA 28552 Db 27452 CATCACCTCATAGAGCCGGGTATTTAACATATGAGGGCCAGGAATTGACTCT 28552 QY 27453 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 2128 CCTGGGAACTGGGCTCTCTAAGCTTACTCTTGTCCGGAAACTCTCTCCAG 2187 Db 27453 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 28512 CTGGGAACTGGGCTACGCGCTCTCAGTTAATCTCTGTGGATGACTGHCCTAGG 28671 Db 27453 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 2188 ATCTGTACTATCTGAGGGGGCTNAAAGGGCGTCACTAGATGATCTTTCAGC 2247 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 28672 ---TCCTTACATCGGAGGAATCTGGGAGCTTCTTCAGTGTGTTTCAAGGTA- -TTTCCACC 28726 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 2248 CACTAAGTATGAGCTGGGAGCTTCTTCAGTGTGTTTCAAGGTA 2307 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 28727 TCCTGTGTGT-AATGGAGACTTCTTCAGTGTGTTTCAAGGTA 28727 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 2308 AGCCCTACTCTGTTAGGGAGGAGCTTACAGAAAGAGGGSCATTACACTG 2367 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 28786 AGTCCCACCTTTATAGGGAGGATAATTAGCCAGGTTGAGCTTATCTACATG 28845 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 2368 AACATAGGAGGAGGAGCCCTTGTGTTGAGGAAATTCATGGTGA 2427 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 28946 AATATGAGAAAGTACCTGAGCTGAGGAGAATTCAGGAGGAACTAACCTGA 28904 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 2428 AGTCTGGCAACGAGACAAATGGAGCGAGCAAAAGAATGCCGTCCTGTCAGGTA 2487 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 28905 AGTCTGGCAATGGAGGAGCAATTTGGAG-GCCAAATACTCCCTGGCCAGTC 28905 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 2488 AACTAAGGAGATCTCTCTTACCAAGGGCTGACCCCTCTGAGCCAAAGGCC 2547 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 28964 GTTAAAGATCCACACTTTCTTATCAAGGAAATCCCTGGCTGAAGGTC 29023 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 2548 AACAGGATCCAAGGATTTAGGACTTAACCTCAAGGCTGTTAAGC 2607 Db 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 QY 27543 GAAATTGCCCCACCCGGGTACATGCCCTCTGATTAAGGAGATCAG 27602 29024 ATAAGGATACAGAAATTGTTAAACATTGAAAGCTCAAGGCTGTTGAGGAATGCA 29083

US-08-979-847B-88

Qy	3198	ARAGTCGACCTGTCACCGAGCTTAACTTACGTGCAAGGCAC	3157
Db	1065	AAAGGCTGGCTCTCTCAGGAGATAGAATCTAGCTAATTCAGGAC	1124
Qy	3158	CARGGCTCTAGGAGGAAYRATCCAGCCTATCTGGCTTATGCCCT	3217
Db	1125	AAAGCAGCTAAGGRGTTCTGGCTTAAGGYTCTGGCAAATGGATTCCCAGGT	3277
Db	1185	AAAGCAGCTAAGGGTTCTGGCTAACGGTTATCCAGCTATGCCCT	1184
Qy	3228	WTGGCGAAATAGCCAGGYCATWATACSTAATTAAGRAACTCGAGGCTAAC	3337
Db	1244	ACASCCOATAGCCAGACCATATACTAACTAATTAGGAACCTCGAGGCTAAC	1303
Qy	3338	ATTATAGTGGAYAMCTGAGMTRAAGTGGCTTCAGGCCCTAAGAACGCC	3397
Db	1304	ATTATAGTAAATGGACCTTACGAGTGGCTTCAGGCCCTAAGAACGCC	1358
Qy	3398	AACCCGAGYCCAGTTAAAGTGGCCAACRGGGCAAGCTTSTTYAIRTCAAGA	3457
Db	1419	-AAAACAGGAAATAGCTAGGCTCTAGGAGCTCTAGGAGTTCTATAGGCCAAGA	1418
Qy	3458	AAAACAGGAACTCTGGAGGCTTACACGRTCCRAGGAGATTTTGACCCR	3517
Db	1419	-AAAACAGGAAATAGCTAGGCTCTAGGAGCTCTAGGAGTTCTATAGGCCAAGA	1477
Qy	3518	TGGCRVACTGASTAGGAAYTGTGTTAGTGGCAGGAGGTGCTYCATTTGAGG	3577
Db	1538	TRATGGNGGCGTAGCAGCTNAGHTCTGAGGKGTTAAATAACGGGAGAGTC	1597
Qy	3638	TPIACTGTCGACATCTCATGAGKPTGAAVRGCATACCTACTGCTAAAGGAGACTTGTGG	1537
Db	1598	TCTCGAGCACCGTACTAAARTCAGGCTTATACTGTAARGGCCAGTGTGCRAC	3757
Qy	3698	TCTCGAGCACCGTACTAAARTCAGGCTTATACTGTAARGGCCAGTGTGCRAC	1657
Db	1658	TGTGACAGCATCTCATGAGKPTGAAVRGCATACCTACTGCTAAAGGAGACTTGTGG	1717
Qy	3758	TGTGACTCTGTGCAACTCTAACCGAGYCNATTCTCCAGACATGAGAAAGATAR	3817
Db	1718	TGGCACTTGTCACCTTAACCT-----	1742
Qy	3818	AAVATTAATCTGTCACCAATTCTAACCTTACGAGGGACCTTGTGARG	3877
Db	1743	-----CAAACTATGTGCCAGAGGAACCTTNTAGAG	1776
Qy	3878	TTCCYVGTACTGATCCYGAACCTTCAACT---TGATATCTGAGGAGTCCTTGAGA	3934
Db	1777	GRCCTTATGCCAACCTGACCTCAACTATATATATACTGATGGAGTGTGTTGAGA	1836
Qy	3935	AAAGGCTCGAAAGGGGATGCC---AGGGTCACTGATAATGGATAATGAACT	3991
Db	1837	AAAGGCTTACAGGAGATATTCAGGAGCTCAACTATATATATACTGATGGAGTGTGTTGAGA	1896
Qy	3992	AAATCCCT---CACTCAGGAACTAGTGTGAGTGTGAGACTAAATAGCCYTCAYT	4045
Db	1897	AGGCCTCTCCCCCAGGGACCGGCCCCGTTAGCAGAAGCTAGTGCACTGACCCG	1956
Qy	4046	KGGGCTCTAGGAGGAGAAAGGGYAAATATATACAGACTCTRATATGT	4105
Db	1957	CGAGCTTAGACTTGTGAGGAGGAGGATATAATGTGATACAGTAGCAACTATGT	2016

14

RESULT 13
 US-08-691-563C-57
 Sequence 57, Application US/08691563C | 102(e)
 Patent No. 6001987
 GENERAL INFORMATION:
 APPLICANT: Herve PERON
 APPLICANT: Frederic BESME
 APPLICANT: Frederic BEIN
 APPLICANT: Glauzia PARAHOS-BACCALA
 APPLICANT: Florence KOMURIAN-PRADEL
 APPLICANT: Colette JOLIVET
 APPLICANT: Bernard MANDRAN
 TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC OR THERAPEUTIC PURPOSES
 NUMBER OF SEQUENCES: 92
 CORRESPONDENCE ADDRESS:
 ADDRESS: Oliff & Berridge
 STREET: 700 South Washington Street, Suite 300
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22314
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/691,563C
 FILING DATE: 02-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Berridge, William P.
 REGISTRATION NUMBER: 30,024
 REFERENCE/DOCKET NUMBER: WPB 38588
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-336-6400
 TELEFAX: 703-836-2787
 INFORMATION FOR SEQ ID NO: 57:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2391 base pairs
 TYPE: nucleotide
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-691-563C-57

Query Match 18.7%; Score 1415.8; DB 3; Length 2391;
 Best Local Similarity 78.4%; Pred. No. 0;
 Matches 1817; Conservative 109; Mismatches 267; Indels 126; Gaps 19;

Qy 2021 ATGATCCAGCGCAGGACTGAGGGTGCCTGGGCCAACGCCAATCCATCACCTC 2080
 Db 1 ATGATCCAGCGCAGGACTGAGGGTGCCTGGGCCAACGCCAATCCATCACCTC 60
 Qy 2081 ACAGAGCCCTGGTAGCTGACCATGTGAGGGCAGGAGGT--TGTCTCTGGACACT 2137

Qy 4105 YACCTAGTCATGCCATGMRCGAATTGTSARAGAAGGGATTCCTRACTCYGAG 4165
 Db 2017 TATCTAATCGAAATGCCCATGTCATAATGAAAGAAGGGAGTTCTTAACCTCTGG 2076
 Qy 4166 RGAAACCTTCATCACGGAGCCATTAGGAAATTATAYGCCTACGAACT 4225
 Db 2077 GGAAACCCCATTAATACCAAGCTTAATGAGTTATGTCACACAGTGCAAACCTC 2136
 Qy 4226 ARAGAGTGAGAGTCTACACTGCGGGCTCATCANAAGGAAGRAAGGAAATASIA 4285
 Db 2137 AAGGGAGGAGCTTACACTGCCAAACCATCAGAAGGGAAAGGGAGGAGGAGCAG 2196
 Qy 4286 GRGAYTGCCTA 4297
 Db 2197 CATAAGTGGCTA 2208

RESULT 14
 US-08-691-563C-57
 Sequence 57, Application US/08691563C | 102(e)
 Patent No. 6001987
 GENERAL INFORMATION:
 APPLICANT: Herve PERON
 APPLICANT: Frederic BESME
 APPLICANT: Frederic BEIN
 APPLICANT: Glauzia PARAHOS-BACCALA
 APPLICANT: Florence KOMURIAN-PRADEL
 APPLICANT: Colette JOLIVET
 APPLICANT: Bernard MANDRAN
 TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC OR THERAPEUTIC PURPOSES
 NUMBER OF SEQUENCES: 92
 CORRESPONDENCE ADDRESS:
 ADDRESS: Oliff & Berridge
 STREET: 700 South Washington Street, Suite 300
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22314
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/691,563C
 FILING DATE: 02-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Berridge, William P.
 REGISTRATION NUMBER: 30,024
 REFERENCE/DOCKET NUMBER: WPB 38588
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-336-6400
 TELEFAX: 703-836-2787
 INFORMATION FOR SEQ ID NO: 57:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2391 base pairs
 TYPE: nucleotide
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-691-563C-57

Query Match 18.7%; Score 1415.8; DB 3; Length 2391;
 Best Local Similarity 78.4%; Pred. No. 0;
 Matches 1817; Conservative 109; Mismatches 267; Indels 126; Gaps 19;

Qy 61 ACAGAGCCCAAGTATCTGACCATGAGGGTCAAGAGGGTACTGTCTCGAGACT 120
 Db 2138 GGCGCGCTCTCTAGCTTACTCTTGTCCCGGAAACTGCTCCAGATCTGACT 2197
 Qy 121 GGCGCGCTCTCTAGCTTACTCTTGTCCCGGAAACTGCTCCAGATCTGACT 180
 Db 2198 ATTCCTGGGGCTCTAGCTAGATCTTTCAGCCACTAAGTTA 2257
 Qy 181 GTCC--GAGGGCTCTAGGAGACGCACTACTGATAC-TTCCTCCAGCACTAAGTIC 236
 Db 2258 TGAATGGGAGCTTTCATGCTTTCATTAGCTGAAAGCCACTA 2317
 Qy 237 TG-ACTGGGACTTACTCTTCCACATGCTTTCATTATGCTGAAGGCCACTC 295
 Db 2318 CCTGTGTAGGAGAGACATCTAGCAAAAGCAGGGCAATTACCTGACAATAGGAG 2377
 Qy 296 TCTGTGGGAGAGACATCTAGCAAAAGCAGGGCAATTACATGAAATAGGAG 355
 Db 2378 AASGAAACCCGTTGTTGTCCTGAGGAGGATATCTGAAAGTGTGGCA 2437
 Qy 356 AAGGAAACATGTGTTGT-CCCTGCTTGAGGAGGAAATCTGAGTGGCA 414
 Db 2438 ACAGAGGCAATGGAGGCAAGAAATGGCTGCTGCTGTCAGTAAGTAACTAAGGA 2497
 Qy 415 ACAGAGGAACTATGAGCAAG-CAGAATGAGTGGCCGCTGTGTCAGTAACTAAGGA 473
 Db 2498 TTCAACTCTTCTCCATCCAAGGAGTACCCCTAGACCAAGGCCAACAGGATT 2557
 Qy 474 TTCCACTCTTCTCCATCCAAGGAGTACCCCTAGACCCAGGCCAACAGGATT 533
 Db 2558 CCAAAGAAGTGTAGGACTTAAGGCCAAGGTTAGTAACCATGATAACTCCCTG 2617
 Qy 534 CCAGAGATGTAAGGACCTAAAGCCAAAGGCTAGTAACCAAGGAAATACCCCTG 593
 Db 2618 CAGTAACTGAGTGTAGGAGGACAGAAACCCGTGGAGCAGTGGAGGTTAGTG 2677
 Qy 594 CAAGACTC-----AATTTAGGATAGGAAACCCACGGAGGTGA-GGTAGTG 645
 Db 2678 CAAGATCTCAGGATTATCATGGGGCGTGTCTTATACCCAGCTGACTAGGCC 2737
 Qy 646 CAAGACTCAGGATTATCAT-GAGGCTGTGTCTATACCCAGCTGTACCA 704
 Db 2738 TTATCTGTGMYTCCAAATACCAAGGAGGAGCAGGAGGTTAACSTCTGGACCTMA 2797
 Qy 705 TTATCTGAGCTCTTCCCHAATACCGAGGAGGAGCAGGAGGTTACGCTGGACCTAA 764
 Db 2798 GGATCCCTTCTGCACTCCGTACACTCTGACTCTCTCAATTCTGTTGCCTGAGA 2857
 Qy 765 GGATCCCTTCTGCACTCCGTACACTCTGACTCTCTCAATTCTGTTGCCTGAGA 824
 Db 2858 TACTCTAACCCCARCACTCACTCACTCCAGCTCTTACCCAGGGTCAGGGATAG 2917
 Qy 825 TCCCTTGAAACCAACTCTCACTCCAGGGTCAAGTGTGTTACCCAGGGTTAGGATAG 884
 Db 2918 YCCCATCTTCTGCACTCCAGGGTCAAGCTGAGCTGAGYCACTTCATCTGGACAC 2977
 Qy 885 CCCCTCTTATTCGGCAGGCTTACGTTAGTACTTGTGACCTCAATTCTCATCTGGACAC 944
 Db 2978 TCTGTGCTCTGAGTGTAGGTTAGTACTTGTGAGCTGAGCTGACATCTGGACAC 3037
 Qy 945 TCTGTGCTCTGAGTGTAGGTTAGTACTTGTGAGCTGACATCTGGACAC 1004
 Db 3038 TAAGGCCAACGAGCTCTMAATTCTCGYACTCTGGCTACAWGGTTCCAAACCS 3097
 Qy 1005 TCAAGCACCAGAACGACTCTTAACCTCTCTACTACTCTGGCTACAGGTTCAC 1064
 Db 3098 ARAGCTCAGCTCTCTAGAGGGTTAACTCTAGGCTTAACTTCAAGGCCAC 3157
 Qy 1065 AAAGGCTCCTCTCTAGGAGGTTACCTGGCTAAATTATTCACAGGCAC 1124
 Qy 3158 CARGGCCCTGAGGAATYATCCAGCTTACTGGCTTACCTCCTACATGCCAAACCT 3217

Db	: : : : : : : : : AACCCAAAGCCCCAGTGTCACTGCTGCCAACAGGCAGATTTCCTTATGCCACAG-	1417
QY	: : : : : : : : AAAACAGRAAYAGCTTGGACTCTAACAGRTCCRAGGAGGGCTTCAACCYR	3517
Db	: : : : : : : AAGAACACAGGAATAGCTCTAGGACTCTAACAGGATGAGCTGACCG	1477
QY	: : : : : TGGCRYACCTGASTAAGGAAAYTGATGTTAGTGGCAAAGGGTGCYCATGTTAVGG	3577
Db	: : : : TGGTATACCTGAGTAGGAATGATGAGTTAGGGCAAGGGTGGCTCAGTGTAGG	1537
QY	: : : TACTGGCGCAAGTAGCAGTYKTAGTATCTGAAAGCTTAAPATACTGGGRAGAGTC	3637
Db	: : TAATGGNGGAGTAGCAGCTNAGTATCTGAAAGCTTAATACGGGAGAGTC	1597
QY	: TACCTGTTGACCTCTGTAAGCTGAACTGTTACCTGCTTAAGGAGCTTGGC	3697
Db	: TNTCTGTGTGACATCTGTGAGTGAARGCATACTCTGCTTAAGGAGCTTGGC	1657
QY	: TGTCAAGACACACYGTACTTAAATCTAGGCTTATTACTGAAARGGCAGTGTGRC	3757
Db	: TGTCAAGACACCAATTACTTAATVATCAGGCTTATTACTGAAAGGCACTGNGAC	1717
QY	: TGTGCACTTGTGRACTCTTAACCCAGYCNCATTCTCCAGACAATSGAGAAAGATAR	3817
Db	: TGGCACTTGTGCACTTAACCC	1742
QY	: 3918 AAVATAACTGCAACAARTAATTCTAACACCTATGCCACTCGAGGGACTCTYTAGARG	3877
Db	: 1743 ----- -----CRAACTTATGCTGCCAGAGGATCTTNTAGAG	1776
QY	: 3878 TCCYTGTGACTGATCCYGACTTCACT--TGATACCTGATGGAGTCCTTGTAGA	3934
Db	: 1777 GTCCCTTAGCCACCTCTGACCTCAACTTATATATACCTGATGAAAGTGTGTTGAGA	1836
QY	: 3935 AAAGGACTTGGAAAGYGGGTATGC--AGTGGTCAGTATAATGGATAYTGAAGT	3991
Db	: 1837 AAGGGATTCAAAGGGNAGGATATCCATGGCTTGTGTTAGTATAAGCAGTACTGAAAGT	1896
QY	: 3992 AATCCCT--CACTCCAGGAACTAGTGTCTYAGCTRGCAACTATACCCY--CAVT	4045
Db	: 1897 AAGCTCTTCCCCCAGGCCACAGGCCCTTGTAGCAGACTAGTGGCACTAACCG	1956
QY	: 4046 KGGGCACTTGAATTAGGAGAAGRAAAAGGYAMATATATACAGACTCTRATATGT	4105
Db	: 1957 CGAACCTTGAACTTGTGAGGGAGGAGATAATGTCTATACAGATACGAACTATGCT	2016
QY	: 4106 YACCTAGTCNTCC----- -----ATGCCATGVRCAATG 4138	
Db	: 2017 TATCTAATCGAAATGCCCTTGTGTTATCTAATCGAAATGCCATGTGCAATAG	2076
QY	: 4139 ARAGAAAGGGAACTCTAACCTCTGAGGAACTATCACATACAGGAGCATTAG	4198
Db	: 2077 AAGGAAGGGAGTCTAACCTCTGGGGACCCCATTAATACACAGTATCATG	2136
QY	: 4199 ARATTTATYGGCTACAGAACCTTCTGAGGAGGGMCTCTACACTGCGGGTCAT	4258
Db	: 2137 GAGTTATGCACTGAGGAAACTCAAGGAGGGAGCTTACACGCAAAAGCCAT	2196
QY	: 4259 CANAAGGAGAAGRAAGGAAATASAAGGAAATGCGAA 4297	
Db	: 2197 CAGAAAGGGAAAGGGAGGAGGAGCATTAAGTGTGCTA 2235	